AMERICAN
JOURNAL OF MATHEMATICS

FOUNDED BY THE JOHNS HOPKINS UNIVERSITY

PUBLISHED UNDER THE JOINT AUSPICES OF

THE JOHNS HOPKINS UNIVERSITY
AND
THE AMERICAN MATHEMATICAL SOCIETY

Volume LXXVIII, Number 2
APRIL, 1956

THE JOHNS HOPKINS PRESS
BALTIMORE 18, MARYLAND
U. S. A.

CONTENTS

The AMERICAN JOURNAL OF MATHEMATICS appears four times yearly.

The subscription price of the JOURNAL is \$8.50 in the U. S.; \$8.75 in Canada; and \$9.00 in other foreign countries. The price of single numbers is \$2.50.

Manuscipts intended for publication in the JOURNAL should be sent to Professor AUREL WINTNER, The Johns Hopkins University, Baltimore 18, Md.

Subscriptions to the JOURNAL and all business communications should be sent to THE JOHNS HOPKINS PRESS, BALTIMORE 18, MARYLAND, U. S. A.

THE JOHNS HOPKINS PRESS supplies to the authors 100 free reprints of every article appearing in the AMERICAN JOURNAL OF MATHEMATICS. On the other hand, neither THE JOHNS HOPKINS PRESS nor the AMERICAN JOURNAL OF MATHEMATICS can accept orders for additional reprints. Authors interested in securing more than 100 reprints are advised to make arrangements directly with the printers, J. H. FURST CO., 20 HOPKINS PLACE, BALTIMORE 1, MARYLAND.

The typescripts submitted can be in English, French, German or Italian and should be prepared in accordance with the instructions listed on the inside back cover of this issue.

Entered as second-class matter at the Baltimore, Maryland, Postoffice, acceptance for mailing at special rate of postage provided for in Section 1103, Act of October 3, 1917, Authorized on July 3, 1918.

PRINTED IN THE UNITED STATES OF AMERICA
BY J. H. FURST COMPANY, BALTIMORE, MARYLAND

SUBALGEBRAS OF THE ALGEBRA OF ALL COMPLEX-VALUED CONTINUOUS FUNCTIONS ON THE CIRCLE.*

By JOHN WERMER.[1]

1. Introduction. Let $\boldsymbol{R}$ be the algebra of all real valued continuous functions on the circle and let $\boldsymbol{C}$ be the algebra of all complex-valued continuous functions on the circle.

The subalgebras we are concerned with are assumed to contain the constant 1.

A fundamental theorem of Stone yields that any uniformly closed subalgebra $\boldsymbol{R}'$ of $\boldsymbol{R}$ which separates points (i.e which is such that if $\lambda_1 \neq \lambda_2$ then there exists an f in $\boldsymbol{R}'$ with $f(\lambda_1) \neq f(\lambda_2)$) coincides with $\boldsymbol{R}$.

For the algebra $\boldsymbol{C}$ the situation is quite different. There exists a large class of proper subalgebras of $\boldsymbol{C}$ which separate points. The problem of classifying these subalgebras leads to the following question: What are the *maximal* subalgebras of $\boldsymbol{C}$?

A closed proper subalgebra $\boldsymbol{M}$ *of* $\boldsymbol{C}$ separating points is called *maximal* if there exists no closed subalgebra $\boldsymbol{M}'$ with $\boldsymbol{M} \subseteq \boldsymbol{M}'$ and $\boldsymbol{M}' \neq \boldsymbol{M}$, $\boldsymbol{M}' \neq \boldsymbol{C}$.

In [1] and [2] the author has given examples of certain maximal subalgebras. Here we shall exhibit a large class of maximal subalgebras, associated with Riemann surfaces.

Let $\mathfrak{F}$ be a Riemann surface,[2] $\mathfrak{M}$ a region on $\mathfrak{F}$ bounded by a simple closed analytic curve γ, such that $\mathfrak{M} + \gamma$ is compact. $\mathfrak{M}$ then has finite genus p. Since γ is topologically a circle, we may regard $\boldsymbol{C}$ as the space of continuous complex-valued functions on γ. We shall use $\| f \|$ to denote $\max\limits_{\lambda \varepsilon \gamma} | f(\lambda) |$.

Definition 1. $\mathfrak{A}$ is the subalgebra of $\boldsymbol{C}$ consisting of all f which may be continued into $\mathfrak{M}$ to be analytic on $\mathfrak{M}$ and continuous on $\mathfrak{M} + \gamma$.

By the maximum principle for $\mathfrak{M}$, $\mathfrak{A}$ is a closed subalgebra of $\boldsymbol{C}$. Also

* Received September 10, 1954.

[1] I am indebted to Professor M. Heins for a number of valuable suggestions which I have used in some of the proofs.

[2] For definitions and basic facts of the theory of Riemann surfaces we refer the reader to R. Nevanlinna, "Uniformisierung," Springer Verlag, 1953.

$\mathfrak{A}$ separates points on γ, since, given p, q on γ, $p \neq q$, we can find some f in $\mathfrak{A}$ with $f(p)=0$, $f(q) \neq 0$.

The main object of this paper is to prove:

THEOREM 2. *$\mathfrak{A}$ is a maximal subalgebra of $\boldsymbol{C}$.*

When $\mathfrak{F}$ is the plane and γ is the unit circle $|\lambda|=1$, $\mathfrak{A}$ becomes the algebra generated by the functions 1 and λ. This is the case discussed in [1].

In Section 2 we prove Theorem 1 in which we give the form of the general linear functional on $\boldsymbol{C}$ which annihilates $\mathfrak{A}$. In Section 3 we use Theorem 1 to prove Theorem 2. In Section 4 we find when two of our maximal subalgebras are isomorphic.

2. Fix ζ in $\mathfrak{M}$; let G_ζ denote the Green's function for $\mathfrak{M}$ singular at ζ. Then G_ζ is harmonic in $\mathfrak{M}$ except at ζ; G_ζ vanishes on γ; for some fixed local parameter z at ζ, $G_\zeta(z)+\log|z-\zeta|$ is regular neighborhood of ζ.

Let H_ζ be the (multiple-valued) conjugate function of G_ζ. Since γ is an analytic curve, $G_\zeta+iH_\zeta$ is analytic everywhere on γ.

Set $W_\zeta(z)=-d\{G_\zeta(z)+iH_\zeta(z)\}/dz$. Then W_ζ is a "covariant" on $\mathfrak{M}$, as defined in [3], p. 102. W_ζ is analytic on $\mathfrak{M}$ except for a simple pole at ζ, with residue 1. On γ we denote by ω_ζ the measure $\frac{1}{2\pi i} W_\zeta(\lambda)d\lambda$. Then ω_ζ is the harmonic measure for $\mathfrak{M}$ evaluated at ζ. In particular, for any set E on γ, $\omega_\zeta(E)$ is real and non-negative.

Fix some ζ_0 in $\mathfrak{M}$. From now on we shall omit the subscript ζ_0 when writing G_{ζ_0}, H_{ζ_0}, W_{ζ_0} or ω_{ζ_0}. We note that W has no zero on γ.

LEMMA 1. *W has $2p$ zeros in $\mathfrak{M}$, where each zero is counted with its multiplicity.*

Proof. We choose an analytic parametrization of $\gamma : \lambda=\lambda(t)$, $0 \leqq t \leqq 1$. In any coordinate neighborhood U of a point of γ with local parameter z, we define $z(t)=z(\lambda(t))$. We can then consider $W(z(t)) \cdot z'(t)$ in $U \cap \gamma$, where the prime indicates differentiation with respect to t. Direct computation shows that this expression is independent of the choice of local parameter. Then

$$\Delta=\int_0^1 d/dt\{\log W(z(t))z'(t)\}dt$$

is well-defined. By a formula given in [3], p. 133, we have

$$\Delta=2\pi i(B-A-N),$$

where B is the number of zeros of W in $\mathfrak{M}$, A the number of poles, and N the Euler characteristic of $\mathfrak{M}$. On the other hand

$$\Delta = \int_0^1 d/dt \log d/dt - \{G(z(t)) + iH(z(t))\}dt$$

and this equals the variation of $\log d - \{G(z(t)) + iH(z(t))\}/dt$ over $0 \leqq t \leqq 1$. The properties of G and H on γ then yield directly that $\Delta = 0$. Thus $B = A + N$. But $A = 1$ and $N = 2p - 1$. Hence $B = 2p$, as asserted.

Let $\alpha_1, \cdots, \alpha_{2p}$ be a homology basis of closed curves on $\mathfrak{M}$.

Definition 2. Let $u(\lambda)$ be a real continuous function on γ and let $U(\zeta)$ be the harmonic function on $\mathfrak{M}$ with $U \equiv u$ on γ. Then $\Phi_\nu(u)$, $\nu = 1, 2, \cdots, 2p$, denotes the period of the conjugate function of U corresponding to α_ν.

Definition 3. For $\mu = 1, \cdots, 2p$, $\psi_\mu(\zeta)$ is a harmonic function on $\mathfrak{M}$, continuous on $\mathfrak{M} + \gamma$ and twice differentiable on γ, with $\Phi_\nu(\psi_\mu) = \delta_\mu{}^\nu$.

LEMMA 2. *There exist functions K_i, $i = 1, \cdots, 2p$, meromorphic on $\mathfrak{M}$ and continuous and real valued on γ such that for $i = 1, \cdots, 2p$, $\Phi_i = K_i d\omega$ as functionals, i.e. $\Phi_i(u) = \int_\gamma u(\lambda) K_i(\lambda) d\omega(\lambda)$, all u, and such that $W(\zeta)K_i(\zeta)$ is analytic on $\mathfrak{M}$ for each i.*

Proof. Let W have the zeros $z_1, \cdots, z_k$ in $\mathfrak{M}$ of orders $\nu_1, \cdots, \nu_k$. By Lemma 1, $\sum_{i=1}^{i=k} \nu_i = 2p$. At z_i we use local polar coordinates (r_i, θ_i) for $i = 1, \cdots, k$.

For each i, $i = 1, \cdots, k$ a well-known construction yields us functions $u_j{}^i$, $j = 1, \cdots, \nu_i$ and $v_j{}^i$, $j = 1, \cdots, \nu_i$ such that:

(i) $u_j{}^i$, $v_j{}^i$ are harmonic on $\mathfrak{M}$ except at z_i.

(ii) $u_j{}^i$ has at z_i a pole with principal part $(r_i)^{-j} \cos j\theta_i$ and $v_j{}^i$ has at z_i a pole with principal part $(r_i)^{-j} \sin j\theta_i{}^j$.

(iii) $u_j{}^i$ and $v_j{}^i$ vanish identically on γ.

The $4p$ functions $u_j{}^i$, $v_j{}^i$ together form a linearly independent set. For suppose $\sum_{i,j} c_j{}^i u_j{}^i + d_j{}^i v_j{}^i = 0$, $c_j{}^i$, $d_j{}^i$ being constants. Because of the poles of the $u_j{}^i$, $v_j{}^i$ this implies $c_j{}^i u_j{}^i + d_j{}^i v_j{}^i = 0$, whence $c_j{}^i \cos j\theta_i + d_j{}^i \sin j\theta_i = 0$. Since θ_i is arbitrary, we conclude $c_j{}^i = d_j{}^i = 0$, all i, j. Thus linear independence is established. We now seek constants $a_j{}^i, b_j{}^i$, $j = 1, \cdots, \nu_i$, $i = 1, \cdots, 2p$,

such that the function $\sum_{i,j} a_j{}^i u_j{}^i + b_j{}^i v_j{}^i$ have a single-valued conjugate function. This gives $2p$ conditions on $4p$ unknowns, and so we obtain at least $2p$ linearly independent $4p$-tuples satisfying the conditions. We thus get the linearly independent functions

$$q^\nu(\zeta) = \sum_{i,j} a_{j\nu}{}^i u_j{}^i(\zeta) + b_{j\nu}{}^i v_j{}^i(\zeta), \qquad \nu = 1, \cdots, 2p.$$

where each q^ν has a single-valued conjugate function. Then $q^\nu(\zeta)$ is harmonic on $\mathfrak{M}$ except for possible poles at the points z_i of orders $\leqq \nu_i$ and $q^\nu(\lambda) = 0$ on γ for all ν.

Let $r^\nu(\zeta)$ be the conjugate function of q^ν with $r^\nu(\zeta_0) = 0$. Then $h_\nu(\zeta) = i(q^\nu(\zeta) + ir^\nu(\zeta))$ is, for $\nu = 1, \cdots, 2p$, a meromorphic function on $\mathfrak{M}$ such that $h_\nu(\lambda)$ is real and continuous on γ for each ν. Suppose $\sum_{\nu=1}^{2p} c_\nu h_\nu = 0$, where $c_\nu = a_\nu + ib_\nu$, $(a_\nu, b_\nu$ real). Since $h_\nu(\lambda)$ is real for λ and γ, this gives $\sum_{\nu=1}^{2p} a_\nu h_\nu(\lambda) = \sum_{\nu=1}^{2p} b_\nu h_\nu(\lambda) = 0$ for λ in γ. But $\sum_{\nu=1}^{2p} a_\nu h_\nu(\zeta)$ is meromorphic and so $\sum_{\nu=1}^{2p} a_\nu h_\nu(\zeta) = 0$, $\zeta \,\varepsilon\, \mathfrak{M}$. Similarly $\sum_{\nu=1}^{2p} b_\nu h_\nu(\zeta) = 0$ on $\mathfrak{M}$. Hence $\sum_{\nu=1}^{2p} a_\nu q^\nu(\zeta) = \sum_{\nu=1}^{2p} b_\nu q^\nu(\zeta) = 0$ on $\mathfrak{M}$. But the q^ν are linearly independent by construction, whence $a_\nu = b_\nu = 0$, all ν. Hence $c_\nu = 0$, all ν. Hence the h_ν are linearly independent.

Consider now the covariant $h_\nu(\zeta) W(\zeta)$. This has no poles except possibly at ζ_0 since the zeros of W cancel the poles of h_ν. Also $h_\nu(\zeta_0) = iq^\nu(\zeta_0)$ and so $\frac{1}{2\pi i}\int_\gamma h_\nu(\lambda) W(\lambda) d\lambda = iq^\nu(\zeta_0)$ by the residue theorem. Now $h_\nu(\lambda)$ is real on γ, and $\frac{1}{2\pi i} W(\lambda) d\lambda$ is a real-valued measure on γ. Hence the left hand side is real. Hence $q^\nu(\zeta_0) = 0$ and so $h_\nu(\zeta_0) = 0$. It follows that $h_\nu(\zeta) W(\zeta)$ is regular at ζ_0 and so everywhere on $\mathfrak{M}$.

Let now N be the space of all real continuous functions u on γ with $\Phi_\nu(u) = 0$ for $\nu = 1, \cdots, 2p$. Given u in N choose v twice differentiable on γ with $\| u - v \| < \epsilon$. Then there is a constant K so that

$$| \Phi_\nu(v) | = | \Phi_\nu(v - u) | \leqq K\epsilon, \qquad \nu = 1, \cdots, 2p.$$

Set $w(\lambda) = v(\lambda) - \sum_{\nu=1}^{2p} \Phi_\nu(v) \psi_\nu(\lambda)$. Then w is differentiable on γ, $\Phi_j(w) = 0$, $j = 1, \cdots, 2p$ and

$$\| u - w \| \leqq \| u - v \| + \sum_{\nu=1}^{2p} |\Phi_\nu(v)| \, \| \psi_\nu \| < \epsilon + K'\epsilon = K''\epsilon, K''$$

independent of ϵ.

Let $w(\zeta)$ be the harmonic function with boundary value $w(\lambda)$. Since $\Phi_j(w)=0$, $j=1,\cdots,2p$, w has a single-valued conjugate w_1 and since w is twice differentiable on γ, w_1 is continuous in $\mathfrak{M}+\gamma$. By the preceding $h_i(\zeta)W(\zeta)$ is analytic on $\mathfrak{M}$ for each i. The residue theorem yields:

$$0=\int_\gamma (w(\lambda)+iw_1(\lambda))h_i(\lambda)W(\lambda)d\lambda,$$

all i. Since $h_i(\lambda)$ is real, we get

$$0=\int_\gamma w(\lambda)h_i(\lambda)W(\lambda)d\lambda=\int_\gamma w(\lambda)h_i(\lambda)dw(\lambda)$$

for all i. Now

$$\left|\int_\gamma u(\lambda)h_i(\lambda)dw(\lambda)\right|=\left|\int_\gamma (u(\lambda)-w(\lambda))h_i(\lambda)dw(\lambda)\right|$$
$$\leqq K''\epsilon\int_\gamma |h_i(\lambda)|\,dw(\lambda).$$

Since ϵ is arbitrary, we get

$$\int_\gamma u(\lambda)h_i(\lambda)dw(\lambda)=0,$$

all i. Thus the functional $h_i dw$ annihilates N. Hence, by elementary vector-space reasoning, there exist constants $b_\nu{}^i$, $i=1,\cdots,2p$, $\nu=1,\cdots,2p$, with

$$h_i(\lambda)dw(\lambda)=\sum_{\nu=1}^{2p} b_\nu\Phi_\nu, \qquad i=1,\cdots,2p.$$

Since the h_i are linearly independent, we can solve this system of equations to get

$$\Phi_\nu=\sum_{i=1}^{2p} c_\nu{}^i h_i(\lambda)dw(\lambda)=K_\nu(\lambda)dw(\lambda), \qquad \nu=1,\cdots,2p.$$

The properties of the h_i established above yield that the K_i satisfy the assertions of the Lemma.

LEMMA 3. *Let μ be any complex-valued Borel measure on γ such that $\int_\gamma f(\lambda)d\mu(\lambda)=0$ whenever $f\in\mathfrak{A}$. Then for closed sets E on γ, $\omega(E)=0$ implies $\mu(E)=0$.*

The analogous assertion was proved for the unit circle by F. and M. Riesz in [4]. A slight modification of their argument yields the following proof.

Proof of Lemma 3. Since E is closed, the complement of E on γ is the union of countably many disjoint arcs γ_n. Since

$$\int_\gamma d\omega(\lambda) < \infty, \qquad \sum_{n=1}^{\infty} \int_{\gamma_n} d\omega(\lambda) < \infty.$$

Hence we can find a sequence of positive numbers d_n with d_n increasing to infinity with n, such that $\sum_{n=1}^{\infty} (\int_{\gamma_n} d\omega(\lambda)) \cdot d_n < \infty$. For $n = 1, 2, \cdots$ we define a positive real twice differentiable function g_n on γ_n such that $I_n = \int_{\gamma_n} g_n(\lambda) d\omega(\lambda) < \infty$ and $g_n(\lambda)$ increases on γ_n to ∞ as λ approaches the endpoints of γ_n. Choose positive constants c_n with $\sum_{n=1}^{\infty} c_n I_n < \infty$.

Set $P(\lambda) = c_n g_n(\lambda) + d_n$ for $\lambda \varepsilon \gamma_n$, $n = 1, 2, \cdots$. Since $\omega(E) = 0$, P is defined almost everywhere on γ with respect to ω and so with respect to ω_ζ for every ζ in $\mathfrak{M}$. Let now $P(\zeta) = \int_\gamma P(\lambda) d\omega_\zeta(\lambda)$, $\zeta \varepsilon \mathfrak{M}$. Note that

$$\int_\gamma P(\lambda) d\omega(\lambda) = \sum_{n=1}^{\infty} (c_n I_n + d_n \int_{\gamma_n} d\omega(\lambda)) < \infty.$$

From the way $P(\lambda)$ was constructed, we see that $P(\lambda)$ is continuous and finite at each point $\lambda \not\varepsilon E$ and $P(\lambda)$ becomes continuously $+\infty$ at each point of E.

By elementary properties of the harmonic measures ω_ζ, we get then that $P(\zeta)$ is harmonic in $\mathfrak{M}$ and has $P(\lambda)$ as continuous boundary function on γ, if we make the obvious definition of continuous approach to ∞ for $P(\zeta)$ as $\zeta \to \lambda$, λ in E. Let $k_1, \cdots, k_{2p}$ be the periods of the conjugate function of $P(\zeta)$. Choose a constant d so that $P_1 = d + P - \sum_{i=1}^{2n} k_i \psi_i > 0$ on γ, and hence on $\mathfrak{M}$ and let Q_1 be the (single-valued) conjugate function of P_1. Set now $k(\zeta) = (1 + P_1 + iQ_1)^{-1} \cdot (P_1 + iQ_1)(\zeta)$. Then $k(\zeta)$ is analytic in $\mathfrak{M}$. For $\lambda \varepsilon \gamma$ and $\lambda \not\varepsilon E$, $k(\lambda) = (1 + x + iy)^{-1} \cdot (x + iy)$ with $0 < x < \infty$ and hence $|k(\lambda)| < 1$. Let $\lambda \varepsilon E$. As $\zeta \to \lambda$ $|P_1(\zeta) + iQ_1(\zeta)| \to \infty$, whence $k(\zeta) \to 1$. Thus $k(\lambda) = 1$ on E. In particular $k(\lambda)$ is continuous in $\mathfrak{M} + \gamma$, and so $k \varepsilon \mathfrak{A}$. Hence $k^n \varepsilon \mathfrak{A}$ for $n = 1, 2, \cdots$. Then

$$0 = \int_\gamma k^n(\lambda) d\mu(\lambda) = \int_E d\mu(\lambda) + \int_{\gamma - E} k^n(\lambda) d\mu(\lambda).$$

Letting $n \to \infty$ and recalling that $|k(\lambda)| < 1$ if $\lambda \varepsilon \gamma - E$, we conclude that $0 = \int_E d\mu(\lambda) = \mu(E)$, as asserted.

Definition 4. We denote by $L^p(\gamma)$ the class of functions $F(\lambda)$ on γ measurable with respect to ω and with $\int_\gamma |F(\lambda)|^p d\omega(\lambda) < \infty$.

COROLLARY. *If μ is a Borel-measure on γ such that $\int_\gamma f(\lambda)d\mu(\lambda) = 0$ for all $f \varepsilon \mathfrak{A}$, then there exists $F(\lambda) \varepsilon L^1(\gamma)$ such that $d\mu(\lambda) = F(\lambda)d\omega(\lambda)$ as measures on γ.*

Proof. By Lemma 3, $\omega(E) = 0$ implies $\mu(E) = 0$ for any closed set E. It follows that this implication holds for each Borel set E.

We can write $\mu = \mu^+ - \mu^- + i\nu^+ - i\nu^-$ where μ^+, μ^-, ν^+, ν^- are real non-negative measures. Let now E be any Borel set with $\omega(E) = 0$. For every Borel subset E' of E, $\omega(E') = 0$ and so $\mu(E') = 0$. Hence $\mu^+(E) = \mu^-(E) = \nu^+(E) = \nu^-(E) = 0$. Thus μ^+, etc. are all absolutely continuous with respect to ω. It follows by the Radon-Nikodym theorem, that $d\mu^+(\lambda) = F_1(\lambda)d\omega(\lambda)$, where $F_1 \varepsilon L^1(\gamma)$. Similiarly $d\mu^-(\lambda) = F_2(\lambda)d\omega(\lambda)$, $F_2 \varepsilon L^1(\gamma)$ and so on. Adding these equations we get the assertion.

Let now γ' be a simple closed analytic curve in $\mathfrak{M}$ such that γ and γ' together bound an annular subregion $\mathfrak{M}'$ of $\mathfrak{M}$. We choose γ' so that all zeros of W lie outside $\mathfrak{M}' + \gamma'$. We can then map $\mathfrak{M}'$ conformally onto the annulus $r' < |z| < 1$ in the plane, by a mapping $z = \chi(\zeta)$, $\zeta \varepsilon \mathfrak{M}'$. Since γ, γ' are analytic curves, χ is analytic on the boundary curves γ and γ'. It follows that for a fixed K, and each Borel set E on γ,

$$\frac{1}{K}\omega(E) \leqq m(\chi(E)) \leqq K\omega(E),$$

where m denotes Lebesgue measure on $|z| = 1$.

Let $F(\zeta)$ be analytic on $\mathfrak{M}'$. Then $F^o(z) = F(\chi^{-1}(z))$ is analytic in the annulus $r' < |z| < 1$. We shall omit the symbol "^{o}," since this omission introduces no ambiguity. Also for $g(\lambda)$ defined on γ, we write $g(e^{i\theta})$ for $g(\chi^{-1}(e^{i\theta}))$.

Definition 5. Let F be analytic in $\mathfrak{M}$. We say $F \varepsilon \mathfrak{H}'$, provided that (with the notations just given)

$$\int_0^{2\pi} |F(re^{i\theta})|\, d\theta = O(1) \text{ as } r \to 1.$$

LEMMA 4. *Let $F(\zeta) \varepsilon \mathfrak{H}'$. Then there exists a function $F^*(\lambda)$ defined on γ* a. e.—*$d\omega$ such that*

(a) $F^* \varepsilon L^1(\gamma)$.

(b) $F(\zeta) = \frac{1}{2\pi i} \int_\gamma F^*(\lambda) W_\zeta(\lambda) d\lambda$, *all* $\zeta \varepsilon \mathfrak{M}$.

(c) $\lim_{\zeta \to \lambda} F(\zeta) = F^*(\lambda)$ a. e.—$d\omega$ *on* γ, *if* $\zeta \to \lambda$ *within some sector.*

(d) *Fix* $r_1 > r'$. *Then for some constant* K *independent of* F *and* r *we have for* $r_1 < r < 1$:

$$\int_0^{2\pi} |F(re^{i\theta})| \, d\theta \leqq K \int_0^{2\pi} |F^*(e^{i\theta})| \, d\theta.$$

Proof. By hypothesis $\int_0^{2\pi} |F(re^{i\theta})| \, d\theta = O(1)$ as $r \to 1$, and $F(z)$ is analytic for $r' < |z| < 1$. We may write $F = F_1 + F_2$ where F_1 is analytic in $|z| < 1$ and F_2 is analytic in $|z| > r'$. We hence get

$$\int_0^{2\pi} |F_1(re^{i\theta})| \, d\theta = O(1) \text{ as } r \to 1.$$

Classical results now give that $F_1^*(e^{i\theta}) = \lim_{z \to e^{i\theta}} F_1(z)$ exists for a. a. θ if the approach to the boundary lies within some sector, and that $F_1^*(e^{i\theta})$ is summable on $0 \leqq \theta < 2\pi$. Hence $\lim_{z \to e^{i\theta}} F(z)$ exists a. e. We denote it by $F^*(e^{i\theta})$; for λ on γ we write $F^*(\lambda)$ instead of $F^*(\chi(\lambda))$. Since sets of ω-measure 0 on γ correspond to sets of Lebesgue measure 0 on $|z| = 1$, we so get assertion (c).

Let now $r' < r < 1$ and let γ_r be the curve in $\mathfrak{M}'$ which χ maps into the circle $|z| = r$. The residue theorem gives for $\zeta \varepsilon \mathfrak{M}$, ζ outside the region bounded by γ_r and γ:

$$F(\zeta) = \frac{1}{2\pi i} \int_{\gamma_r} F(\lambda) W_\zeta(\lambda) d\lambda = \frac{1}{2\pi} \int_0^{2\pi} F(re^{i\theta}) W_\zeta(re^{i\theta}) re^{i\theta} d\theta.$$

By a classical theorem, $\lim_{r \to 1} \int_0^{2\pi} |F_1(re^{i\theta}) - F_1^*(e^{i\theta})| \, d\theta = 0$; also $W_\zeta(re^{i\theta})$ is continuous for $r' < r \leqq 1$. It follows that

$$F(\zeta) = \lim_{r \to 1} \frac{1}{2\pi} \int_0^{2\pi} F(re^{i\theta}) W_\zeta(re^{i\theta}) re^{i\theta} \, d\theta = \frac{1}{2\pi} \int_0^{2\pi} F^*(e^{i\theta}) W_\zeta(e^{i\theta}) e^{i\theta} \, d\theta.$$

Hence $F(\zeta) = \frac{1}{2\pi i} \int_\gamma F^*(\lambda) W_\zeta(\lambda) d\lambda$, $\zeta \varepsilon \mathfrak{M}$. Thus (b) is proved.

Now

$$\int_\gamma |F^*(\lambda)| \, d\omega(\lambda) = \frac{1}{2\pi} \int_0^{2\pi} |F^*(e^{i\theta})| \, W(e^{i\theta}) e^{i\theta} \, d\theta < \infty,$$

since $\int_0^{2\pi} |F^*(e^{i\theta})|\, d\theta < \infty$ by construction of F^*, and W is continuous. Thus (a) holds. Finally, write $F(z) = F_1(z) + F_2(z)$, where $F_1(z)$ is analytic in $|z| < 1$ and $F_2(z) = -\frac{1}{2\pi i}\int_{|\tau|=r'} F(\tau)(\tau - z)^{-1} d\tau$; so that F_2 is analytic in $|z| > r'$. Let $\zeta = \chi^{-1}(\tau)$. Then $F(\tau) = \int_\gamma F^*(\lambda)(W(\lambda))^{-1} W_\zeta(\lambda) d\omega(\lambda)$ by (b), whence $|F(\tau)| \leqq M_\zeta \int_\gamma |F^*(\lambda)|\, d\omega(\lambda)$, where $M_\zeta = \max_{\lambda \varepsilon \gamma} |W^{-1}(\lambda) W_\zeta(\lambda)|$. Hence for $0 \leqq \theta < 2\pi$, $r' < r$,

$$|F_2(re^{i\theta})| \leqq 1/(r - r') \max_{|\tau|=r'} |F(\tau)| \leqq (M/2\pi) 1/(r - r') \int_\gamma |F^*(\lambda)|\, d\omega(\lambda),$$

where $M = \sup_{\zeta \varepsilon \gamma'} M_\zeta$. At last, $r' < r_1 \leqq r$,

$$\int_0^{2\pi} |F_2(re^{i\theta})|\, d\theta \leqq \int_\gamma |F^*(\lambda)|\, d\omega(\lambda) \cdot M_1 \leqq M_2 \cdot \int_0^{2\pi} |F^*(e^{i\theta})|\, d\theta,$$

where M_1 and M_2 are constants. On the other hand, since F_1 is analytic in $|z| < 1$,

$$\int_0^{2\pi} |F_1(re^{i\theta})|\, d\theta \leqq \int_0^{2\pi} |F_1^*(e^{i\theta})|\, d\theta \leqq \int_0^{2\pi} |F^*(e^{i\theta})|\, d\theta + \int_0^{2\pi} |F_2^*(e^{i\theta})|\, d\theta$$
$$\leqq (1 + M_2) \int_0^{2\pi} |F^*(e^{i\theta})|\, d\theta.$$

Hence

$$\int_0^{2\pi} |F(re^{i\theta})|\, d\theta = \int_0^{2\pi} |F_1(re^{i\theta}) + F_2(re^{i\theta})|\, d\theta$$
$$\leqq (1 + 2M_2) \int_0^{2\pi} |F^*(e^{i\theta})|\, d\theta.$$

This proves (d).

Definition 6. $\mathfrak{B}$ is the conjugate space of $\boldsymbol{C}$.

By the representation theorem of F. Riesz, $\mathfrak{B}$ may be identified with the space of all complex-valued Borel-measures on γ.

Definition 7. $\mathfrak{W}$ is the subspace of $\mathfrak{B}$ consisting of all measures μ of the form $d\mu(\lambda) = G^*(\lambda) d\omega(\lambda)$ where G^* is the boundary function of some G in $\mathfrak{H}'$ with $G(\zeta_0) = 0$.

LEMMA 5.[3] *$\mathfrak{W}$ is regularly closed as subspace of $\mathfrak{B}$.*

[3] The idea of using a lemma of this kind resulted from a conversation with Professor S. Kakutani.

Proof. By a theorem of Banach, [5], p. 124, it suffices to show that with each weakly convergent sequence of elements of $\mathfrak{W}$ the limit again is in $\mathfrak{W}$.

Let now $\mu_n \varepsilon \mathfrak{W}$, μ_n converge weakly to μ. By Definition 7, there exists $G_n \varepsilon \mathfrak{H}'$, $G_n(\zeta_0) = 0$ with $d\mu_n(\lambda) = G_n^*(\lambda) d\omega(\lambda)$. Then for each $f \varepsilon \mathfrak{A}$, $f \cdot G_n \varepsilon \mathfrak{H}'$, and so by Lemma 4,

$$0 = f(\zeta_0) G_n(\zeta_0) = \int_\gamma f(\lambda) G_n^*(\lambda) d\omega(\lambda) = \int_\gamma f(\lambda) d\mu_n(\lambda).$$

For $f \varepsilon \mathfrak{A}$, then, $0 = \int_\gamma f(\lambda) d\mu(\lambda)$. By the Corollary to Lemma 3, this implies that there exists $G_0 \varepsilon L^1(\gamma)$ with $G_0(\lambda) d\omega(\lambda) = d\mu(\lambda)$.

We shall show that G_n converges to a function G analytic on $\mathfrak{M}$ with $G \varepsilon \mathfrak{H}'$ and that $G^*(\lambda) = G_0(\lambda)$ a. e.—$d\omega$ on γ. From this it follows that $d\mu(\lambda) = G^*(\lambda) d\omega(\lambda)$ and so $\mu \varepsilon \mathfrak{W}$. Now

$$G_n(\zeta) = \frac{1}{2\pi i} \int_\gamma G_n^*(\lambda) W_\zeta(\lambda) d\lambda = \int_\gamma (W(\lambda))^{-1} W_\zeta(\lambda) d\mu_n(\lambda).$$

Hence $G(\zeta) = \lim_{n\to\infty} G_n(\zeta)$ exists for $\zeta \varepsilon \mathfrak{M}$. Also

$$| G_n(\zeta) | \leqq M_\zeta \int_\gamma | G_n^*(\lambda) | d\omega(\lambda),$$

where $M_\zeta = \max_{\lambda \varepsilon \gamma} | (W(\lambda))^{-1} W_\zeta(\lambda) |$. Now since the sequence μ_n converges weakly, the total variation of μ_n has a bound K valid for all n. Hence

$$\int_\gamma | G_n^*(\lambda) | d\omega(\lambda) < K. \tag{1}$$

Also M_ζ is bounded on each compact subset of $\mathfrak{M}$. Hence by Vitali's theorem, G is analytic on $\mathfrak{M}$ and $\lim_{n\to\infty} G_n = G$ uniformly on each compact subset of $\mathfrak{M}$. Now

$$\int_\gamma | G_n^*(\lambda) | d\omega(\lambda) = \frac{1}{2\pi} \int_0^{2\pi} | G_n^*(e^{i\theta}) | W(e^{i\theta}) e^{i\theta} d\theta.$$

But $W(e^{i\theta}) e^{i\theta}$ has a positive lower bound on $(0, 2\pi)$. Hence $\int_0^{2\pi} | G_n^*(e^{i\theta}) | d\theta < K'$, all n, by (1). Hence $\int_0^{2\pi} | G_n(re^{i\theta}) | d\theta < K''$, all n, by (d) of Lemma 4, $r' < r_1 < r < 1$. It follows that $\int_0^{2\pi} | G(re^{i\theta}) | d\theta < K''$, whence $G \varepsilon \mathfrak{H}'$. We claim that $G^*(\lambda) = G_0(\lambda)$ a. e.—$d\omega$.

Let now $r' < r < 1$. Set

$$U_n^1(z) = \frac{1}{2\pi i}\int_{|\tau|=r'} (\tau - z)^{-1} G_n(\tau)\, d\tau, \quad |z| > r'$$

$$U_n^2(z) = \frac{1}{2\pi i}\int_{|\tau|=r} (\tau - z)^{-1} G_n(\tau)\, d\tau, \quad |z| < r.$$

Then U_n^1 is analytic for $|z| > r'$, U_n^2 for $|z| < 1$, and

$$G_n(z) = U_n^2(z) - U_n^1(z), \quad r' < |z| < 1.$$

Clearly $U_n^{1*}(e^{i\theta}) = \lim\limits_{z\to e^{i\theta}} U_n^1(z)$ exists for all θ. Also $G_n^*(e^{i\theta})$ exists a. e. Hence $U_n^{2*}(e^{i\theta})$ exists a. e. and $G_n^* = U_n^{2*} - U_n^{1*}$, a. e.

Now $G_n(\tau) \to G(\tau)$ uniformly on $|\tau| = r'$. Hence $U_n^1(z) \to U^1(z)$ uniformly for $r' < a \leqq |z| \leqq b < \infty$, and $U^{1*}(e^{i\theta})$ exists everywhere. Hence $U_n^2(z) \to U^2(z)$ uniformly in $r' < a \leqq |z| \leqq b < 1$, and $G(z) = U^2(z) - U^1(z)$. It follows that

$$\int_0^{2\pi} |U^2(re^{i\theta})|\, d\theta = O(1) \text{ as } r \to 1. \tag{2}$$

Also

$$G^*(e^{i\theta}) = U^{2*}(e^{i\theta}) - U^{1*}(e^{i\theta}) \text{ a. e.} \tag{3}$$

Fix r, ϕ; $r < 1$, $0 \leqq \phi < 2\pi$. Set $g(\theta) = (1 - r^2)(1 + r^2 - 2r\cos(\theta - \phi))^{-1}$. Now

$$\frac{1}{2\pi}\int_0^{2\pi} U_n^{2*}(e^{i\theta}) g(\theta)\, d\theta = \frac{1}{2\pi}\int_0^{2\pi} (G_n^*(e^{i\theta}) + U_n^{1*}(e^{i\theta})) g(\theta)\, d\theta$$

and $U_n^{1*}(e^{i\theta}) \to U^{1*}(e^{i\theta})$ uniformly in $0 \leqq \theta < 2\pi$ and $G_n^*(\lambda) d\omega(\lambda)$ converges weakly to $G_0(\lambda) d\omega(\lambda)$. Hence

$$\lim_{n\to\infty} \frac{1}{2\pi}\int_0^{2\pi} U_n^{2*}(e^{i\theta}) g(\theta)\, d\theta = \frac{1}{2\pi}\int_0^{2\pi} (G_0(e^{i\theta}) + U^{1*}(e^{i\theta})) g(\theta)\, d\theta.$$

On the other hand,

$$U_n^2(re^{i\phi}) = \frac{1}{2\pi}\int_0^{2\pi} U_n^{2*}(e^{i\theta}) g(\theta)\, d\theta \to U^2(re^{i\phi}).$$

By (2) we get, since U^2 is analytic in $|z| < 1$,

$$U^2(re^{i\phi}) = \frac{1}{2\pi}\int_0^{2\pi} U^{2*}(e^{i\theta}) g(\theta)\, d\theta.$$

Hence

$$\int_0^{2\pi} U^{2*}(e^{i\theta}) g(\theta)\, d\theta = \int_0^{2\pi} (G_0(e^{i\theta}) + U^1(e^{i\theta})) g(\theta)\, d\theta,$$

or

$$0=\int_0^{2\pi}\{U^{2*}-G_0-U^{1*}\}(e^{i\theta})(1-r^2)(1+r^2-2r\cos(\theta-\phi))^{-1}d\theta.$$

This now holds for arbitrary r, ϕ. Since $U^{2*}-G_0-U^{1*}$ is summable on $(0,2\pi)$, we conclude

$$G_0(e^{i\theta})=U^{2*}(e^{i\theta})-U^{1*}(e^{i\theta})=G^*(e^{i\theta}),\ \text{a.e.}$$

The conclusion now follows, as shown above.

THEOREM 1. *Let* $\mu_0 \varepsilon \mathfrak{B}$ *and* $\int_\gamma f(\lambda)d\mu_0(\lambda)=0$ *for all* f *in* $\mathfrak{A}$. *Then there exists* $J \varepsilon \mathfrak{H}'$, $J(\zeta_0)=0$, *and constants* c_i *such that, setting*

$$L(\lambda)=J^*(\lambda)+\sum_{i=1}^{2p}c_iK_i(\lambda),$$

where K_i *are the functions of Lemma 2, we have* $d\mu_0(\lambda)=L(\lambda)d\omega(\lambda)$.

Proof. Let $\mathfrak{W}'$ be the vector-space obtained by adjoining to $\mathfrak{W}$ the measures $K_i(\lambda)d\omega(\lambda)$, $i=1,\cdots,2p$. Since $\mathfrak{W}$ is regularly closed, the same is true of $\mathfrak{W}'$. Our assertion amounts to the statement that $\mu_0 \varepsilon \mathfrak{W}'$.

Suppose $\mu_0 \not\varepsilon \mathfrak{W}'$. Since $\mathfrak{W}'$ is regularly closed, it follows by Banach's definition, that for some f_0 in $\boldsymbol{C}$

$$\int_\gamma f_0(\lambda)d\mu_0(\lambda)\neq 0 \tag{4}$$

$$\int_\gamma f_0(\lambda)d\mu(\lambda)=0 \text{ if } \mu\,\varepsilon\,\mathfrak{W}'. \tag{5}$$

Let now $\bar{\mathfrak{A}}$ be the closure in $L^2(\gamma)$ of $\mathfrak{A}$. Then we can decompose f_0 as follows: $f_0=H+G$, $H \varepsilon \bar{\mathfrak{A}}$, G orthogonal to $\bar{\mathfrak{A}}$. Let $f \varepsilon \mathfrak{A}$. Then

$$\int_\gamma \bar{f}(\lambda)G(\lambda)d\omega(\lambda)=0. \tag{6}$$

Hence

$$\int_\gamma G(\lambda)d\omega(\lambda)=0. \tag{6'}$$

Let now $H_n \varepsilon \mathfrak{A}$, $H_n \to H$ in the norm of $L^2(\gamma)$. Then by the residue theorem

$$\frac{1}{2\pi i}\int_\gamma H_n(\lambda)(f(\lambda)-f(\zeta_0))d\omega(\lambda)=0,$$

whence

$$\int_\gamma H(\lambda)(f(\lambda)-f(\zeta_0))d\omega(\lambda)=0. \tag{7}$$

Also $(f(\lambda)-f(\zeta_0))d\omega(\lambda)\,\varepsilon\,\mathfrak{W}'$, whence by (5)

$$\int_\gamma f_0(\lambda)(f(\lambda)-f(\zeta_0))d\omega(\lambda)=0.$$

Hence from $f_0=H+G$ and (7),

$$\int_\gamma G(\lambda)(f(\lambda)-f(\zeta_0))d\omega(\lambda)=0.$$

By (6′), then,

(8) $$\int_\gamma G(\lambda)f(\lambda)d\omega(\lambda)=0.$$

It follows from (6) and (8) that $\int_\gamma u(\lambda)G(\lambda)d\omega(\lambda)=0$ for all real continuous functions u on γ with $u=\operatorname{Re}f$, for some $f\,\varepsilon\,\mathfrak{A}$. As in the proof of Lemma 2, we get from this that for some b_ν, $G(\lambda)d\omega(\lambda)=\sum_{\nu=1}^{2p}b_\nu\Phi_\nu$ as functionals, and hence that

(9) $$G(\lambda)d\omega(\lambda)=\sum_{i=1}^{2p}c_iK_i(\lambda)d\omega(\lambda)$$

where K_i are the functions constructed in Lemma 2. Now

$$\int_\gamma H(\lambda)K_\nu(\lambda)d\omega(\lambda)=0, \qquad \nu=1,\cdots,2p;$$

also $K_\nu(\lambda)d\omega(\lambda)\,\varepsilon\,\mathfrak{W}'$. Hence by (5),

$$\int_\gamma f_0(\lambda)K_\nu(\lambda)d\omega(\lambda)=0, \qquad \nu=1,\cdots,2p.$$

Hence

(10) $$\int_\gamma G(\lambda)K_\nu(\lambda)d\omega(\lambda)=0, \qquad \nu=1,\cdots,2p.$$

By (9) and (10),

$$\int_\gamma |G(\lambda)|^2d\omega(\lambda)=\int_\gamma G(\lambda)\sum_{i=1}^{2p}\bar{c}_iK_i(\lambda)d\omega(\lambda)$$
$$=\sum_{i=1}^{2p}\bar{c}_i\int_\gamma G(\lambda)K_i(\lambda)d\omega(\lambda)=0.$$

Hence $G(\lambda)=0$ a.e. and so $f_0=H$ a.e.

Now consider $H_n\,\varepsilon\,\mathfrak{A}$, $H_n\to H$ in the norm of $L^2(\gamma)$. Then for $\zeta\,\varepsilon\,\mathfrak{M}$,

$$H_n(\zeta)=\frac{1}{2\pi i}\int_\gamma H_n(\lambda)W_\zeta(\lambda)d\lambda.$$

It then follows that

$$\frac{1}{2\pi i}\int_\gamma f_0(\lambda)W_\zeta(\lambda)d\lambda = \lim_{n\to\infty} H_n(\zeta)$$

is analytic in $\mathfrak{M}$. Also $\frac{1}{2\pi i}\int_\gamma f_0(\lambda)W_\zeta(\lambda)d\lambda$ has f_0 as continuous boundary value on γ. Hence $f_0 \varepsilon \mathfrak{A}$. Then $\int_\gamma f_0(\lambda)d\mu_0(\lambda) = 0$. This contradicts (4). Hence the assertion must be true.

3. Proof of Theorem 2. Let $\mathfrak{A}'$ be a closed subalgebra of $\boldsymbol{C}$ with $\mathfrak{A}' \neq \boldsymbol{C}$ and $\mathfrak{A} \subseteq \mathfrak{A}'$. Since $\mathfrak{A}'$ is a proper closed subspace of $\boldsymbol{C}$, a well-known theorem on Banach spaces guarantees the existence of a non-zero functional on $\boldsymbol{C}$ which annihilates $\mathfrak{A}'$. Thus there exists $\mu \varepsilon \mathfrak{B}$, $\mu \neq 0$ with $\int_\gamma g(\lambda)d\mu(\lambda) = 0$ if $g \varepsilon \mathfrak{A}'$. This holds in particular if $g \varepsilon \mathfrak{A}$. Hence by Theorem 1, $d\mu(\lambda) = L_0(\lambda)d\omega(\lambda)$ where L_0 is meromorphic on $\mathfrak{M}$ and analytic on $\mathfrak{M}$ except at the poles of the K_i. Hence L_0 is analytic except at the points $z_1, \cdots, z_k$ where W vanishes. Also, $\lim_{\zeta\to\lambda} L_0(\zeta)$ exists for a. a. λ on γ, if $\zeta \to \lambda$ within some sector, and this limit $\neq 0$ a. e. on γ.

Fix now $\phi \varepsilon \mathfrak{A}'$. We shall show $\phi \varepsilon \mathfrak{A}$. For if $f \varepsilon \mathfrak{A}$, $f(\lambda)\phi^m(\lambda) \varepsilon \mathfrak{A}'$ for $m = 1, 2, \cdots$ whence $\int_\gamma f(\lambda)\phi^m(\lambda)d\mu(\lambda) = 0$. Applying Theorem 1 to the measures $\phi^m(\lambda)d\mu(\lambda)$, we get $\phi^m(\lambda)d\mu(\lambda) = L_m(\lambda)d\omega(\lambda)$ where L_m has the same analyticity and boundary behavior as L_0. Hence $\phi^m(\lambda)L_0(\lambda) = L_m(\lambda)$ a. e. on γ. It follows that $(L_1(\lambda))^m = L_m(\lambda)(L_0(\lambda))^{m-1}$ a. e. on γ. On both sides we have non-tangential boundary values of functions analytic in the region $\mathfrak{M}_0$ obtained by deleting from $\mathfrak{M}$ the points $z_1, \cdots, z_k$. By a result of Lusin and Privaloff, [6], an analytic function possessing non-tangential boundary values on a set of positive measure is determined by these values. Hence $(L_1(\zeta))^m = L_m(\zeta)(L_0(\zeta))^{m-1}$ for ζ in $\mathfrak{M}_0$. Since this is true for all $m \geqq 1$, L_0 cannot have a zero at any point ζ' in $\mathfrak{M}_0$ of order α unless L_1 has at ζ' a zero of order $\geqq \alpha$. Hence $L_0^{-1}L_1$ is analytic in $\mathfrak{M}_0$. Also, since $\phi(\lambda)L_0(\lambda) = L_1(\lambda)$ a. e. on γ, ϕ is the non-tangential limit of $L_0^{-1}L_1$ a. e. on γ.

Set $T\phi(\zeta) = L_0^{-1}(\zeta)L_1(\zeta)$. The map $\phi \to T\phi$ then assigns to each ϕ in $\mathfrak{A}'$ an analytic function $T\phi$ on $\mathfrak{M}_0$ having boundary values $\phi(\lambda)$. By the theorem in [6] mentioned above, ϕ determines $T\phi$. Let now ϕ_1, ϕ_2 belong to $\mathfrak{A}'$. Then

$$\operatorname*{Lim}_{\zeta\to\lambda} T\phi_1(\zeta) \cdot T\phi_2(\zeta) = \phi_1(\lambda)\phi_2(\lambda)$$

and so $T(\phi_1 \cdot \phi_2) = T\phi_1 \cdot T\phi_2$. Similarly $T(\phi_1 + \phi_2) = T\phi_1 + T\phi_2$. Fix now z_0 in $\mathfrak{M}_0$. Then the map $\phi \to T\phi(z_0)$ is a multiplicative functional defined on $\mathfrak{A}'$. But a multiplicative functional on a Banach algebra is always bounded and has bound 1. Hence $|T\phi(z_0)| \leqq \|\phi\|$. Since z_0 is an arbitrary point in $\mathfrak{M}_0$, $T\phi$ is then bounded on $\mathfrak{M}_0$; hence $T\phi$ is analytic and bounded on $\mathfrak{M}$. Lemma 4 gives now that for ζ in $\mathfrak{M}$,

$$T\phi(\zeta) = \int_\gamma (T\phi)^*(\lambda)\, d\omega_\zeta(\lambda) = \int_\gamma \phi(\lambda)\, d\omega_\zeta(\lambda).$$

On the other hand the last integral represents a continuous function on $\mathfrak{M} + \gamma$ agreeing with $\phi(\lambda)$ on γ. Hence ϕ is in $\mathfrak{A}$, as asserted.

Hence $\mathfrak{A}' = \mathfrak{A}$, and so Theorem 2 is established.

4. (Added November 27, 1954.) Let now $\mathfrak{F}_1$, $\mathfrak{F}_2$ be Riemann surfaces, $\mathfrak{M}_1$, $\mathfrak{M}_2$ regions on them bounded by simple closed analytic curves γ_1, γ_2 with $\mathfrak{M}_i \cup \gamma_i$ compact, $i = 1, 2$. Let $\mathfrak{A}_i$ be the algebra of functions continuous on γ_i and extendable to be analytic on $\mathfrak{M}_i$, $i = 1, 2$. We assert:

THEOREM 3. *$\mathfrak{A}_1$ is isomorphic to $\mathfrak{A}_2$ as algebra if and only if $\mathfrak{M}_1$ is conformally equivalent to $\mathfrak{M}_2$.*

We need the following:

LEMMA. *If χ is a multiplicative functional on $\mathfrak{A}_i$, then there exists a point $p \varepsilon \mathfrak{M}_i \cup \gamma_i$ with $\chi(f) = f(p)$, all $f \varepsilon \mathfrak{A}_i$.*

Proof.[4] (We omit the subscript i from $\mathfrak{A}_i$, etc.) By the general representation theorem for bounded linear functions on spaces of continuous functions, there is a measure μ_0 on γ with

$$\chi(f) = \int_\gamma f(\lambda)\, d\mu_0(\lambda), \qquad f \varepsilon \mathfrak{A}.$$

Suppose now that the assertion of the Lemma is false. Then for each $p \varepsilon \mathfrak{M} \cup \gamma$ there exists $f_p \varepsilon \mathfrak{A}$ with $\chi(f_p) = 0$ and $f_p(p) \neq 0$.

Let $d\omega(\lambda)$ and $W(\zeta)$ have the same meaning as in the preceding sections. Let $\mathfrak{M}_0$ be the region obtained by deleting from $\mathfrak{M}$ the zeros of W.

Now for all $f \varepsilon \mathfrak{A}$ and p in $\mathfrak{M} \cup \gamma$

$$0 = \chi(f \cdot f_p) = \int_\gamma f(\lambda) f_p(\lambda)\, d\mu_0(\lambda).$$

Hence the measure $f_p(\lambda)\, d\mu_0(\lambda)$ annihilates $\mathfrak{A}$. By Theorem 1, then we can

[4] Cf. L. Carleson [7], Theorem 4, for a similar method of proof.

find a function L_p analytic on $\mathfrak{M}_0$ and with $L_p(\zeta)W(\zeta)$ regular on $\mathfrak{M}$, such that L_p has nontangential boundary-values $L_p(\lambda)$ for all λ in γ except for a set of ω-measure 0, and with $f_p(\lambda)d\mu_0(\lambda)=L_p(\lambda)d\omega(\lambda)$ as measures. Choose now p_1, p_2 distinct in $\mathfrak{M}\cup\gamma$. Then

$$f_{p_2}(\lambda)f_{p_1}(\lambda)d\mu_0(\lambda)=f_{p_2}(\lambda)L_{p_1}(\lambda)d\omega(\lambda)$$

and

$$f_{p_1}(\lambda)f_{p_2}(\lambda)d\mu_0(\lambda)=f_{p_1}(\lambda)L_{p_2}(\lambda)d\omega(\lambda)$$

whence $f_{p_2}\cdot L_{p_1}=f_{p_1}\cdot L_{p_2}$ a. e.—$d\omega$ on γ, whence by the result in [6] which we have quoted earlier, $f_{p_2}(\zeta)L_{p_1}(\zeta)=f_{p_1}(\zeta)L_{p_2}(\zeta)$ for all ζ in $\mathfrak{M}$.

Fix now p_0 in $\mathfrak{M}$ and set $F(\zeta)=f_{p_0}^{-1}(\zeta)L_{p_0}(\zeta)$. Since, for $q\,\varepsilon\,\mathfrak{M}_0$, L_q *and* f_q^{-1} are regular at q, we obtain that F is regular at q. Thus F is analytic on all $\mathfrak{M}_0$ and similarly we see that the covariant $F(\zeta)W(\zeta)$ is analytic on all of $\mathfrak{M}$.

Next, for each $q\,\varepsilon\,\gamma$, we choose an arc γ_q on γ with $|f_q(\lambda)|\geqq\delta_q$ for λ in γ_q, δ_q being a positive number. By the Heine-Borel theorem, some finite set of these arcs covers γ. We can hence get $\delta>0$ and a decomposition $\gamma=\bigcup_{i=1}^{n}\gamma_i$ where the γ_i are disjoint half-open arcs and for each i there is some q_i with $|f_{q_i}(\lambda)|\geqq\delta$ on γ_i.

Now, $F(\zeta)=f_{q_i}^{-1}(\zeta)L_{q_i}(\zeta)$ for $\zeta\,\varepsilon\,\mathfrak{M}$, whence

$$F(\lambda)=\lim_{\zeta\to\lambda}F(\zeta)=f_{q_i}^{-1}(\lambda)L_{q_i}(\lambda)\text{ a.e. on }\gamma_i.$$

We now use annular coordinates r,θ: $r_0\leqq r\leqq 1$, $0\leqq\theta\leqq 2\pi$ in an annular subregion of $\mathfrak{M}$ bounded on one side by γ, with $r=1$ being the equation of γ.

Let $g\,\varepsilon\,\mathfrak{A}$. Then for each i

$$\int_{\gamma_i}g(\lambda)d\mu_0(\lambda)=\int_{\gamma_i}g(\lambda)f_{q_i}^{-1}(\lambda)L_{q_i}(\lambda)d\omega(\lambda)=\int_{\gamma_i}F(\lambda)g(\lambda)d\omega(\lambda).$$

Hence

$$\int_{\gamma}g(\lambda)d\mu_0(\lambda)=\int_{\gamma}g(\lambda)F(\lambda)(2\pi i)^{-1}W(\lambda)d\lambda.$$

Now if γ_ρ is the curve with equation: $r=\rho$, $\rho<1$,

$$\int_{\gamma_\rho}F(\zeta)g(\zeta)W(\zeta)d\zeta=0$$

by the residue theorem. Also

$$\lim_{\rho\to 1}\int_{\gamma_\rho} F(\zeta)g(\zeta)W(\zeta)d\zeta = \int_{\gamma} g(\lambda)F(\lambda)W(\lambda)d\lambda$$

due to the boundary behavior of the functions L_p and f_p. Hence

$$\chi(g) = \int_{\gamma} g(\lambda)d\mu_0(\lambda) = \frac{1}{2\pi i}\int_{\gamma} g(\lambda)F(\lambda)W(\lambda)d\lambda = 0.$$

This must hold for all $g \varepsilon \mathfrak{A}$, which is impossible. Hence the assertion of the Lemma must be true.

COROLLARY. *The space* $\mathfrak{S}$ *of multiplicative functionals on* $\mathfrak{A}$ *is homeomorphic to the set* $\mathfrak{M} \cup \gamma$.

Proof. By the Lemma, if $\chi \varepsilon \mathfrak{S}$, then there exists $p \varepsilon \mathfrak{M} \cup \gamma$ with $\chi(f) = f(p)$ for all $f \varepsilon \mathfrak{A}$. There cannot exist two distinct points p_1, p_2 with this property, for if $p_1 \neq p_2$ then for some f in $\mathfrak{A}$, $f(p_1) \neq f(p_2)$. Hence the map $\chi \to p$ takes $\mathfrak{S}$ into $\mathfrak{M} \cup \gamma$. It is obviously one-one and it is onto $\mathfrak{M} \cup \gamma$ since each p in $\mathfrak{M} \cup \gamma$ defines some multiplicative functional on $\mathfrak{A}$. Finally, the map is easily seen to be bicontinuous.

Proof of Theorem 3. Let τ be an algebraic isomorphism of $\mathfrak{A}_1$ onto $\mathfrak{A}_2$. Fix p in $\mathfrak{M}_2 \cup \gamma_2$. Map each f in $\mathfrak{A}_1$ into $\tau(f)(p)$. This map is a multiplicative functional on $\mathfrak{A}_1$, whence by the lemma there exists $\phi(p)$ in $\mathfrak{M}_1 \cup \gamma_1$ with $\tau(f)(p) = f(\phi(p))$ if $f \varepsilon \mathfrak{A}_1$. The function ϕ then maps $\mathfrak{M}_2 \cup \gamma_2$ onto $\mathfrak{M}_1 \cup \gamma_1$ in a one-one and bicontinuous fashion. It follows that ϕ maps $\mathfrak{M}_2$ homeomorphically onto $\mathfrak{M}_1$.

Fix p_0 in $\mathfrak{M}_2$ and f_0 in $\mathfrak{A}_1$ with f_0 locally simple at $\phi(p_0)$. Then for p in some neighborhood of p_0, $f_0(\phi(p)) = \tau(f_0)(p)$. Since f_0 and $\tau(f_0)$ are analytic functions and moreover f_0 is one-one in a neighborhood of $\phi(p_0)$, ϕ is analytic at p_0 as mapping from $\mathfrak{M}_2$ to $\mathfrak{M}_1$. This holds for each p_0 in $\mathfrak{M}_2$ and further ϕ is globally one-one. Hence ϕ provides a conformal map of $\mathfrak{M}_2$ onto $\mathfrak{M}_1$.

Conversely, suppose we are given a conformal map ϕ of $\mathfrak{M}_2$ on $\mathfrak{M}_1$. Classical results then give that ϕ is extendable to a homeomorphism of $\mathfrak{M}_2 \cup \gamma_2$ onto $\mathfrak{M}_1 \cup \gamma_1$. For each f in $\mathfrak{A}_1$ we can then define τf on $\mathfrak{M}_2 \cup \gamma_2$ as follows: $\tau f(p) = f(\phi(p))$, $p \varepsilon \mathfrak{M}_2 \cup \gamma_2$. Then $\tau f \varepsilon \mathfrak{A}_2$ and τ is an isomorphism from $\mathfrak{A}_1$ to $\mathfrak{A}_2$. This proves Theorem 3.

BROWN UNIVERSITY.

REFERENCES.

[1] J. Wermer, "On algebras of continuous functions," *Proceedings of the American Mathematical Society*, vol. 4 (1953), pp. 866-869.

[2] ———, "Algebras with two generators," *American Journal of Mathematics*, vol. 76 (1954), pp. 853-859.

[3] R. Nevanlinna, *Uniformisierung*, Springer Verlag, 1953.

[4] F. and M. Riesz, "Über Randwerte einer analytischen Funktion," *Quatrième Congrès des math. scandinaves*, 1916, pp. 27-44.

[5] S. Banach, *Opérations Linéaires*, Warszawa, 1932.

[6] N. N. Lusin and I. I. Privaloff, "Sur l'unicité et multiplicité des fonctions analytiques," *Annales de l'École Normale Supérieure*, ser. 3, vol. 42 (1925), pp. 143-191.

[7] L. Carleson, "On bounded analytic functions and closure problems," *Arkiv för Matematik*, 2, 12 (1952), pp. 283-291.

SOME RESULTS ON COHOMOTOPY GROUPS.*

By Franklin P. Peterson.[1]

1. Introduction. One of the central problems of topology is the computation of the set of homotopy classes of maps of a complex K into the n-sphere S^n. In 1936, Borsuk [2] showed that if the dimension of $K = N \leqq 2n - 2$, then this set admits a natural abelian group structure. In this case, this set is called the n-th cohomotopy group of K and denoted by $\pi^n(K)$. In 1949, Spanier [13] derived the basic properties of these groups and expressed the existing theorems on the structure of $\pi^n(K)$ by means of an exact sequence [13; p. 240]. These theorems are the Hopf theorem [7], which states that the natural homomorphism $\eta^n : \pi^n(K) \to H^n(K)$ ($=$ the n-th cohomology group of K) is an isomorphism for $n = N$ and is onto in case $n = N - 1$, and the Steenrod theorem [14], which computes the kernel of η^{N-1} and the image of η^{N-2}. Little more is known about the structure of $\pi^n(K)$.

In this paper, we shall derive further results concerning the structure of $\pi^n(K)$. First, $\pi^n(K)$ is finitely generated when K is finite. Second, $\pi^n(K)$ and $H^n(K)$ have the same rank. Third, the Hopf result is generalized by determining, for each prime p, a range of values of n for which η^n gives an isomorphism on the p-primary components of $\pi^n(K)$ and $H^n(K)$. Finally, the Steenrod result is generalized by giving a computation of the kernel and cokernel of η^n restricted to the p-primary components for a range of values of n where η^n is not an isomorphism. This computation is given in terms of the reduced p-th power operations of Steenrod [15]. In proving these results, we make use of a cohomotopy exact couple similar to that of Massey [8; part III] and of Serre's technique of "isomorphisms modulo a class of groups" [11].

In conclusion, I wish to express my warm appreciation to Professor N. E. Steenrod for his kind advice and encouragement. This paper is essentially Part I of a paper written under his direction and submitted as a dissertation to Princeton University.

* Received September 16, 1955.

[1] The author was a predoctoral National Science Foundation Fellow during the preparation of this paper.

2. Preliminaries. In this section, we recall the notions and notations which we need in order to state our main results.

We first recall the definition and elementary properties of cohomotopy groups [13]. Let K be a finite dimensional CW-complex [6], and let L be a subcomplex. Let $a:(K,L)\to(S^n,\text{pt.})$ be a continuous map, where S^n denotes the n-dimensional sphere and "pt." denotes any fixed point of S^n. Let $[a]$ denote the homotopy class of a. The set of all such homotopy classes has a natural abelian group structure defined on it if dimension $K=N\leqq 2n-2$. We call this the n-th cohomotopy group of the CW-pair (K,L) and denote it by $\pi^n(K,L)$. A map $f:(K,L)\to(K',L')$ induces a homomorphism $f^{\#}:\pi^n(K',L')\to\pi^n(K,L)$ defined by $f^{\#}([a])=[af]$.

Let $\pi_r(X)$ denote the r-th homotopy group of the space X. The process of suspension induces a homomorphism $S_{\#}:\pi_r(S^n)\to\pi_{r+1}(S^{n+1})$ which is an isomorphism for $r<2n-1$ by the Freudenthan theorem [16]. We identify these groups under this isomorphism and denote the result by $Z_{(r-n)}$.

The homology theory best suited for our investigations is the cellular homology theory as described in [6]. We denote the n-th homology group of (K,L) with coefficients in G by $H_n(K,L;G)$ and the n-th cohomology group of (K,L) with coefficients in G by $H^n(K,L;G)$.

We denote the additive group of integers by Z, the group of integers mod n by Z_n, and the p-primary component of a group A by A_p, where the p-primary component of A is the subgroup of all elements of A whose orders are a power of the prime p. (The only exception to this notation is Z_p which denotes the integers mod p.) Let $\phi:A\to B$ be a homomorphism. We denote $\phi|A_p:A_p\to B_p$ by $\phi_{(p)}$. Furthermore, we denote the kernel of ϕ by $\operatorname{Ker}\phi$, the image of ϕ by $\operatorname{Im}\phi$, and the cokernel of ϕ $(=B/\operatorname{Im}\phi)$ by $\operatorname{Coker}\phi$. $A\otimes B$ and $\operatorname{Tor}(A,B)$ denote the tensor product and the torsion product respectively [4].

We now recall the notion of a class which was introduced by Serre [11]. A *class* $\mathcal{C}$ is a non-empty family of abelian groups such that

(I) if $0\to A\to B\to C\to 0$ is an exact sequence [5], then $A\,\varepsilon\,\mathcal{C}$ and $C\,\varepsilon\,\mathcal{C}$ if and only if $B\,\varepsilon\,\mathcal{C}$.

In the applications, one of the following axioms is also assumed:

(II_A) if $A\,\varepsilon\,\mathcal{C}$ and $B\,\varepsilon\,\mathcal{C}$, then $A\otimes B\,\varepsilon\,\mathcal{C}$ and $\operatorname{Tor}(A,B)\,\varepsilon\,\mathcal{C}$, or

(II_B) if $A\,\varepsilon\,\mathcal{C}$ and B is arbitrary, then $A\otimes B\,\varepsilon\,\mathcal{C}$ and $\operatorname{Tor}(A,B)\,\varepsilon\,\mathcal{C}$.

The important examples of classes are $\mathcal{C}_0=$ the class consisting of the

0 group alone, $\mathcal{C}_T =$ the family of torsion groups, $\mathcal{C}_p =$ the family of torsion groups whose p-primary components are 0, $\mathfrak{F} =$ the family of finitely generated groups, $\mathcal{C}_f =$ the family of finite groups, and $\mathcal{C}_{pf} =$ the family of finite groups whose p-primary components are 0. It is easily checked that $\mathcal{C}_0$, $\mathcal{C}_T$, and $\mathcal{C}_p$ satisfy axiom ($\mathrm{II_B}$), while $\mathfrak{F}$, $\mathcal{C}_f$, and $\mathcal{C}_{pf}$ satisfy axiom ($\mathrm{II_A}$) but not ($\mathrm{II_B}$).

The notion of class was introduced by Serre to allow us to ignore systematically certain groups. With this in mind, we make the following definitions: a homomorphism $\phi: A \to B$ is a *$\mathcal{C}$-monomorphism* if $\operatorname{Ker}\phi \varepsilon \mathcal{C}$; it is a *$\mathcal{C}$-epimorphism* if $\operatorname{Coker}\phi \varepsilon \mathcal{C}$; it is a *$\mathcal{C}$-isomorphism* if both Ker ϕ and Coker $\phi \varepsilon \mathcal{C}$.

For any class $\mathcal{C}$, let $\alpha(\mathcal{C})$ denote the largest integer such that $Z_{(s)} \varepsilon \mathcal{C}$ for $0 < s < \alpha(\mathcal{C})$.

THEOREM 2.1. (a) *$Z_{(r)}$ is finite if $r > 0$,*

(b) $(Z_{(r)})_p = Z_p$ *if* $r = 2p - 3$

$= 0$ *otherwise for* $r < 4p - 5$,

(c) $\alpha(\mathcal{C}_T) = \infty$,

(d) $\alpha(\mathcal{C}_p) = 2p - 3$, *and*

(e) $\alpha(\mathcal{C}_0) = 1$.

Proof. (a) is a result of Serre [10]. (b) is a result of Serre [11]. (c) follows from (a). (d) follows from (a) and (b). It is well-known that $Z_{(1)} = \pi_{n+1}(S^n) = Z_2$, hence $\alpha(\mathcal{C}_0) = 1$.

3. The main results. The purpose of this section is to state our main results on the structure of cohomotopy groups. The remainder of this paper is devoted to the proofs of these results.

There is a natural homomorphism $\eta^r: \pi^r(K, L) \to H^r(K, L)$. η^r is defined as follows: let $a \varepsilon [a] \varepsilon \pi^r(K, L)$, and let u be a chosen generator of $H^r(S^r, \text{pt.})$. Then $\eta^r([a]) = a^*(u) \varepsilon H^r(K, L)$ (see Section 4 and [13; p. 234]). We study the relations between the cohomotopy groups and the cohomology groups using this homomorphism. The classical Hopf theorem [7] states that if K is an N-dimensional complex, then η^N is an isomorphism. Our first result extends this theorem modulo classes.

Let (K, L) be a CW-pair with dimension $K = N$ throughout the rest of this paper.

THEOREM 3.1. *Let $\mathcal{C}$ be a class satisfying condition* ($\mathrm{II_B}$) *of Section 2. Let $n > (N+1)/2$ be such that $H^r(K,L) \varepsilon \mathcal{C}$ for every $r > n$. Then η^r is a $\mathcal{C}$-isomorphism if $r > \mathrm{Max}((N+1)/2, n-\alpha(\mathcal{C}))$, and is a $\mathcal{C}$-epimorphism for $r = n - \alpha(\mathcal{C})$ in case $n - \alpha(\mathcal{C}) > (N+1)/2$.*

THEOREM 3.2. *Let $\mathcal{C}$ be a class satisfying condition* ($\mathrm{II_A}$) *of Section 2. Let $n > \mathrm{Max}((N+1)/2, N-\alpha(\mathcal{C}))$ be such that $H^r(K,L) \varepsilon \mathcal{C}$ for every $r > n$. Then η^r is a $\mathcal{C}$-isomorphism if $r \geqq n$, and is a $\mathcal{C}$-epimorphism for $r = n-1$.*

Theorems 3.1 and 3.2 solve a problem proposed by Steenrod in [9]. We have as an immediate corollary of Theorem 3.1:

COROLLARY 3.3. *Let K and L be two CW-complexes of dimensions M and N respectively, and let $f: L \to K$. Let $\mathcal{C}$ be a class satisfying condition* ($\mathrm{II_B}$) *of Section 2, and let $n > \mathrm{Max}((M+1)/2, (N+2)/2)$. Then the following two statements are equivalent:* (a) $f^*: H^r(K) \to H^r(L)$ *is a $\mathcal{C}$-isomorphism for $r > n$ and a $\mathcal{C}$-epimorphism for $r = n$, and* (b) $f^\#$: $\pi^r(K) \to \pi^r(L)$ *is a $\mathcal{C}$-isomorphism for $r > n$ and a $\mathcal{C}$-epimorphism for $r = n$.*

Proof. Replace f by a cellular approximation f' [6; p. 98]. By the mapping cylinder construction [6; p. 108], we may assume f' is an inclusion. Then (a) is true if and only if $H^r(K,L) \varepsilon \mathcal{C}$ for $r > n$ by the exact cohomology sequence of a pair. This is true if and only if $\pi^r(K,L) \varepsilon \mathcal{C}$ for $r > n$ by Theorem 3.1. However, this is true if and only if (b) is true by the exact cohomotopy sequence of a pair.

By specializing $\mathcal{C}$ to particular classes, we have the following four corollaries of Theorems 3.1 and 3.2.

COROLLARY 3.4. *Let $n > (N+1)/2$ be such that $H^r(K,L)$ is finitely generated for every $r > n$. Then $\pi^r(K,L)$ is finitely generated for $r > n$.*

Proof. This follows immediately from 3.2 by setting $\mathcal{C} = \mathcal{F}$ and noting 2.1 (a).

COROLLARY 3.5. *Let $n > (N+1)/2$ be such that $H^r(K,L)$ is finitely generated for every $r > n$. Then $\pi^r(K,L)$ and $H^r(K,L)$ have the same rank for $r > n$. Furthermore, if $r > (N+1)/2$ and $u \varepsilon H^r(K,L)$, then there is an integer $M \neq 0$ such that $Mu \varepsilon \mathrm{Im}\, \eta^r$.*

Proof. By 3.4, $\pi^r(K,L)$ and $H^r(K,L)$ are finitely generated for every $r > n$. Now apply 3.1 with $n = N$, $\mathcal{C} = \mathcal{C}_T$, and note that $\alpha(\mathcal{C}_T) = \infty$ by 2.1 (c). The conclusion that η^r is a $\mathcal{C}_T$-isomorphism for $r > n$ means that

$\pi^r(K,L)$ and $H^r(K,L)$ have the same rank. Furthermore, if for some $u \in H^r(K,L)$ there did not exist a non-zero integer M such that $Mu \in \operatorname{Im} \eta^r$, then $\operatorname{Coker} \eta^r \not\in \mathcal{C}_T$.

The above is a result of Serre [11; p. 288].

COROLLARY 3.6. *Let $n > (N+1)/2 + 1$ be such that $H^r(K,L) = 0$ for every $r > n$. Then $\pi^r(K,L) = 0$ for $r > n$, η^n is an isomorphism, and η^{n-1} is an epimorphism.*

Proof. Set $\mathcal{C} = \mathcal{C}_0$, and use Theorem 3.1. Note that $\mathcal{C}_0$-isomorphism means regular isomorphism.

COROLLARY 3.7. *Let $n > (N+1)/2$ be such that $H^r(K,L) \in \mathcal{C}_p$ for every $r > n$. Then $\eta^r{}_{(p)}$ is an isomorphism for $r >$ Max$((N+1)/2, n-2p+3)$ and $\eta^{n-2p+3}{}_{(p)}$ is an epimorphism if $n - 2p + 3 > (N+1)/2$.*

Proof. Set $\mathcal{C} = \mathcal{C}_p$, and use Theorem 3.1. Note that ϕ being a $\mathcal{C}_p$-isomorphism implies that $\phi_{(p)}$ is an isomorphism.

For completeness, we state without proof a slight extension of a result of Steenrod [14] and Spanier [13; p. 240]. The theorem follows from an unpublished result of Adem (see [8; p. 263]) in a manner analogous to the way Theorem 3.9 follows from Theorem 6.2. Let $Sq^2 : H^{n-2}(K,L) \to H^n(K,L;Z_2)$ denote the Steenrod square [14]. Let $\Lambda : H^n(K,L;Z_2) \to \eta^{n-1}(K,L)$ be the homomorphism defined by Spanier [13; p. 238].

THEOREM 3.8. *Let $n > (N+1)/2 + 2$ be such that $H^r(K,L) = 0$ for every $r > n$. Then $\pi^r(K,L) = 0$ for $r > n$, η^n is an isomorphism, and the following sequence is exact:*

$$\pi^{n-2}(K,L) \xrightarrow{\eta^{n-2}} H^{n-2}(K,L) \xrightarrow{Sq^2} H^n(K,L;Z_2) \xrightarrow{\Lambda} \pi^{n-1}(K,L)$$
$$\xrightarrow{\eta^{n-1}} H^{n-1}(K,L) \longrightarrow 0.$$

Our final main result extends Corollary 3.7 in a manner analogous to the way Theorem 3.8 extends Corollary 3.6. Let

$$\mathcal{P}^1 : H^s(K,L) \to H^{s+2p-2}(K,L;Z_p)$$

denote the first Steenrod reduced p-th power operation [15].

$$\mu : H^{r+2p-3}(K,L;Z_p) \to \pi^r(K,L)_p$$

is a homomorphism to be defined in Section 6. In Section 6 we prove the following:

THEOREM 3.9. *Let $n > (N+1)/2$ be such that $H^r(K,L) \varepsilon \mathcal{C}_p$ for every $r > n$. Then $\eta^r{}_{(p)}$ is an isomorphism for $r >$* Max$((N+1)/2, n-2p+3)$ *and the following sequences are exact:*

$$H^{r-1}(K,L) \xrightarrow{\mathcal{P}^1} H^{r+2p-3}(K,L;Z_p) \xrightarrow{\mu} \pi^r(K,L)_p \xrightarrow{\eta^r{}_{(p)}} H^r(K,L)_p$$

$$\xrightarrow{\mathcal{P}^1{}_{(p)}} H^{r+2p-2}(K,L;Z_p) \quad \textit{for } r > \mathrm{Max}((N+1)/2, n-4p+5)$$

and

$$\eta^{n-4p+5}(K,L)_p \xrightarrow{\eta^{n-4p+5}{}_{(p)}} H^{n-4p+5}(K,L)_p \xrightarrow{\mathcal{P}^1{}_{(p)}} H^{n-2p+3}(K,L;Z_p)$$

if $n-4p+5 > (N+1)/2+1$.

For $r > n-4p+5$, note that Theorem 3.9 computes the kernel and the cokernel of $\eta^r{}_{(p)}$ in terms of the cohomology groups of (K,L) and the first reduced p-th powers in $H^*(K,L)$.

4. The cohomotopy exact couple. The proofs of our main theorems are based on a cohomotopy exact couple of the pair (K,L) similar to the one studied by Massey [8; part III]. Since it differs from Massey's cohomotopy exact couple, we describe it in detail.

Let (K,L) be a CW-pair with dimension $K = N$. Let z be the least integer $> (N+1)/2$; i.e. z is the least integer n for which $\pi^n(K,L)$ has a natural group structure. Let K^s denote the union of L with the s-dimensional skeleton of K. Our exact couple is based on the exact cohomotopy sequence of the triple (K,K^s,K^{s-1}):

$$\pi^z(K,K^s) \xrightarrow{j} \pi^z(K,K^{s-1}) \to \cdots \to \pi^r(K,K^s) \xrightarrow{j} \pi^r(K,K^{s-1})$$

$$\xrightarrow{i} \pi^r(K^s,K^{s-1}) \xrightarrow{\Delta} \pi^{r+1}(K,K^s) \to \cdots,$$

where i and j are the homomorphisms induced by the inclusions

$$(K^s,K^{s-1}) \to (K,K^{s-1}) \text{ and } (K,K^{s-1}) \to (K,K^s)$$

respectively and Δ is the coboundary operator of the triple (K,K^s,K^{s-1}) (see [13; p. 229] for the definition of Δ and the proof of exactness).

For notational convenience, we set

$$A^{r,s}=\pi^r(K,K^s) \text{ for } r \geqq z,$$

$$C^{r,s}=\pi^r(K^s,K^{s-1}) \text{ for } r \geqq z.$$

Also let

$$j^{r,s}\colon A^{r,s}\to A^{r,s-1},$$

$$i^{r,s}\colon A^{r,s-1}\to C^{r,s}, \text{ and}$$

$$\Delta^{r,s}\colon C^{r,s}\to A^{r+1,s}$$

be the appropriate j, i, or Δ for $r \geqq z$. In order to extend the above sequence to an exact sequence extending indefinitely in both directions, we set

$$C^{z-1,s}=\operatorname{Ker} j^{z,s} \text{ for } r<z-1,$$

$$C^{r,s}=0 \text{ for } r<z-1,$$

$$A^{r,s}=0 \text{ for } r<z,$$

$\Delta^{z-1,s}\colon C^{z-1,s}\to A^{z,s}$ to be the inclusion, and the remaining homomorphisms i, j, and Δ to be zero.

The indices on the homomorphisms i, j, and Δ are determined by their domains, and thus we omit them whenever possible.

The groups and homomorphisms defined above fit together in a lattice as in Figure 1. Any path in Figure 1 which moves downward and to the right in a zig-zag pattern traces out an exact sequence. This follows immediately from the exact sequence of the triple (K,K^s,K^{s-1}) and our definitions.

$$\begin{tikzcd}
 & & \vdots \arrow[d] & & \vdots \arrow[d] & \\
 & \cdots \arrow[r] & A^{r,s} \arrow[r, "i"] \arrow[d, "j"] & C^{r,s+1} \arrow[r, "\Delta"] & A^{r+1,s+1} \arrow[r] \arrow[d, "j"] & \cdots \\
\cdots \arrow[r] & C^{r-1,s-1} \arrow[r, "\Delta"] & A^{r,s-1} \arrow[r, "i"] \arrow[d, "j"] & C^{r,s} \arrow[r, "\Delta"] & A^{r+1,s} \arrow[r] \arrow[d, "j"] & \cdots \\
 & \cdots \arrow[r] & A^{r,s-2} \arrow[r, "i"] \arrow[d] & C^{r,s-1} \arrow[r, "\Delta"] & A^{r+1,s-1} \arrow[r] \arrow[d] & \cdots \\
 & & \vdots & & \vdots &
\end{tikzcd}$$

Figure 1.

We now compute some of the groups of the cohomotopy exact couple. The following groups are obviously 0 for $m \geqq 1$: $C^{r,r-m}$, $C^{N+m,s}$, $A^{N+m,s}$, $A^{r,N+m-1}$, and $C^{r,N+m}$. Also $j\colon A^{r,r-m}\to A^{r,r-m-1}$ is an isomorphism for $m \geqq 2$ and is onto for $m=1$. This follows because

$$C^{r-1,r-m}\to A^{r,r-m}\to A^{r,r-m-1}\to C^{r,r-m}$$

is exact and for $m \geqq 1$, $C^{r,r-m}=0$. Hence $A^{r,r-2}\approx A^{r,-1}=\pi^r(K,L)$.

Define a homomorphism

$$\psi\colon \pi^r(K^s, K^{s-1}) \to C^s(K, L; \pi_s(S^r, \mathrm{pt.}))$$

as follows. The cellular homology and cohomology theory is based on $\pi_s(K^s, K^{s-1})$ as the group of chains in dimension s [7]. Hence

$$C^s(K, L; \pi_s(S^r, \mathrm{pt.})) = \mathrm{Hom}(\pi_s(K^s, K^{s-1}), \pi_s(S^r, \mathrm{pt.})),$$

where $\mathrm{Hom}(A, B)$ denotes the group of all homomorphisms from A to B. Let $[b] \in \pi_s(K^s, K^{s-1})$, $[a] \in \pi^r(K^s, K^{s-1})$, then define

$$\psi([a])([b]) = [ab] \in \pi_s(S^r, \mathrm{pt.}).$$

It is shown in [13; p. 222] that ψ has the following properties for $r \geqq z$ and $s \leqq N$:

1) ψ is an isomorphism,

2) ψ is natural with respect to cellular maps $f\colon (K, L) \to (K', L')$, and

3) the following diagram is commutative:

$$\begin{array}{ccc} \pi^r(K^s, K^{s-1}) & \xrightarrow{i\Delta} & \pi^{r+1}(K^{s+1}, K^s) \\ \downarrow \psi & & \downarrow \psi \\ C^s(K, L; \pi_s(S^r, \mathrm{pt.})) & \xrightarrow{S_\# \delta} & C^{s+1}(K, L; \pi_{s+1}(S^{r+1}, \mathrm{pt.})), \end{array}$$

where δ is the coboundary homomorphism

$$C^s(K, L; \pi_s(S^r, \mathrm{pt.})) \to C^{s+1}(K, L; \pi_s(S^r, \mathrm{pt.}))$$

and $S_\#$ is the homomorphism

$$C^{s+1}(K, L; \pi_s(S^r, \mathrm{pt.})) \to C^{s+1}(K, L; \pi_{s+1}(S^{r+1}, \mathrm{pt.}))$$

induced by suspending the coefficient group. Under our assumptions on r and s, $S_\#$ is an isomorphism on the coefficients, and we may write the commutative diagram as

$$\begin{array}{ccc} \pi^r(K^s, K^{s-1}) & \xrightarrow{i\Delta} & \pi^{r+1}(K^{s+1}, K^s) \\ \downarrow \psi & & \downarrow \psi \\ C^s(K, L; Z_{(s-r)}) & \xrightarrow{\delta} & C^{s+1}(K, L; Z_{(s-r)}) \end{array}$$

(see Section 2 for the definition of $Z_{(s-r)}$).

We now describe the first derived cohomotopy exact couple (see [8; part I] for the prescise definitions and the proof that this is an exact couple). Define

$$\mathcal{H}^{r,s} = H(C^{r,s}) = \mathrm{Ker}(i^{r+1,s+1}\Delta^{r,s})/\mathrm{Im}(i^{r,s}\Delta^{r-1,s-1}),$$

$$\Gamma^{r,s} = \mathrm{Im}\, j^{r,s},$$

$$j'^{r,s}\colon \Gamma^{r,s} \to \Gamma^{r,s-1} \text{ by } j^{r,s-1},$$

$$i'^{r,s}\colon \Gamma^{r,s-1} \to \mathcal{H}^{r,s} \text{ by } i^{r,s}(j^{r,s-1})^{-1}, \text{ and}$$

$$\Delta'^{r,s}\colon \mathcal{H}^{r,s} \to \Gamma^{r+1,s+1} \text{ by } \Delta^{r,s}.$$

These groups and homomorphisms fit together in a lattice as in Figure 2.

From the remarks above, the following groups are obviously 0 for $m \geqq 1$: $\mathcal{H}^{r,r-m}$, $\mathcal{H}^{N+m,s}$, $\Gamma^{N+m,s}$, $\Gamma^{r,N+m-1}$, and $\mathcal{H}^{r,N+m}$. Also

$$\Gamma^{r,r-1} \approx \Gamma^{r,r-2} \approx \cdots \approx \pi^r(K,L) \text{ for } r \geqq z.$$

Furthermore, $\mathcal{H}^{r,s} \approx H^s(K,L;Z_{(s-r)})$ for $r \geqq z+1$ by the above identifications and definitions of $\mathcal{H}^{r,s}$.

$$\begin{tikzcd}
 & & \vdots \arrow[d] & & \vdots \arrow[d] & \\
 & \cdots \arrow[r] & \Gamma^{r,s} \arrow[r, "i'"] \arrow[d, "j'"] & \mathcal{H}^{r,s+1} \arrow[r, "\Delta'"] & \Gamma^{r+1,s+2} \arrow[r] \arrow[d, "j'"] & \cdots \\
\cdots \arrow[r] & \mathcal{H}^{r-1,s-2} \arrow[r, "\Delta'"] & \Gamma^{r,s-1} \arrow[r, "i'"] \arrow[d, "j'"] & \mathcal{H}^{r,s} \arrow[r, "\Delta'"] & \Gamma^{r+1,s+1} \arrow[r] \arrow[d, "j'"] & \cdots \\
 & \cdots \arrow[r] & \Gamma^{r,s-2} \arrow[r, "i'"] \arrow[d] & \mathcal{H}^{r,s-1} \arrow[r, "\Delta'"] & \Gamma^{r+1,s} \arrow[r] \arrow[d] & \cdots \\
 & & \vdots & & \vdots &
\end{tikzcd}$$

Figure 2.

Under these identifications, $i'^{r,r}\colon \pi^r(K,L) \to H^r(K,L)$ is defined geometrically as follows. Let $[a] \varepsilon \pi^r(K,L)$. We may assume that $a\colon (K,L) \to (S^r, \text{pt.})$ is such that $a(K^{r-1}) = \text{pt}$. Restrict a to a map $a'\colon (K^r, K^{r-1}) \to (S^r, \text{pt.})$, then $[a']$ is an r-cochain of (K,L) which is a cocycle. Its cohomology class is $i'^{r,r}([a])$. We denote the homomorphism $i'^{r,r}$ by η^r. It is easily checked that this gives the same definition as given in Section 3.

We recall for reference later the definition of a cohomology operation θ of type $(n,q;A,B)$. θ is a function $\theta\colon H^n(K,L;A) \to H^q(K,L;B)$, defined for every CW-pair (K,L), such that if $f\colon (K,L) \to (K',L')$, then $f^*\theta = \theta f^*\colon H^n(K',L';A) \to H^q(K,L;B)$. A theorem of Serre [12; p. 220] states that the cohomology operations of a given type $(n,q;A,B)$ are in 1-1 correspondence with the elements of $H^q(A,n;B) = H^p(K(A,n);B)$, where $K(A,n)$ is an Eilenberg-MacLane complex (see [12] for details). This theorem holds if θ is only defined for every CW-pair (K,L) with dimension $K \leqq N$ and $N \geqq q+1$.

Let $f\colon (K,L)\to(K',L')$ be a cellular map; i.e. $f(K^n)\subset K'^n$. Then f induces homomorphisms

$$f^{\#}\colon \pi^r(K'^s,K'^{s-1})\to\pi^r(K^s,K^{s-1}) \quad\text{and}\quad f^{\#}\colon \pi^r(K',K'^s)\to\pi^r(K,K^s)$$

which commute with i, j, and Δ when $r\geqq z$, z', where $z'=$ the least integer $>$ (dimension $K'+1)/2$. Hence f induces a homomorphism of the cohomotopy exact couple of (K',L') into the cohomotopy exact couple of (K,L), and hence a homomorphism of the first derived cohomotopy exact couples. Thus this induced homomorphism f^* commutes with $i'\Delta'\colon \mathcal{H}^{r,s}\to\mathcal{H}^{r+1,s+2}$. Hence, $i'\Delta'$ is a cohomology operation because any two cellular approximations f_1 and f_2 to an arbitrary map $g\colon (K,L)\to(K',L')$ induce the same homomorphism g^* on $\mathcal{H}^{r,s}$ and $\mathcal{H}^{r+1,s+2}$ (see [6; p. 98] for the definition and properties of cellular approximations).

5. Proof of the Hopf theorem mod $\mathcal{C}$. In this section we give the proofs of Theorems 3.1 and 3.2.

Proof of 3.1. The proof is based on the first derived cohomotopy exact couple.[2] In order to prove that η^r is a $\mathcal{C}$-isomorphism, it suffices to show that $\Gamma^{r,r}\,\varepsilon\,\mathcal{C}$ for $r>\operatorname{Max}((N+1)/2, n-\alpha(\mathcal{C}))$ because

$$\Gamma^{r,r}\xrightarrow{\ j'\ }\pi^r(K,L)\xrightarrow{\ \eta^r\ }H^r(K,L)\xrightarrow{\ \Delta'\ }\Gamma^{r+1,r+1}$$

is an exact sequence. Again by exactness (and the fact that $\Gamma^{r,N}=0\,\varepsilon\,\mathcal{C}$), it suffices to prove that $H^{r+1}(K,L;Z_{(1)})\,\varepsilon\,\mathcal{C},\cdots,H^N(K,L;Z_{(N-r)})\,\varepsilon\,\mathcal{C}$ for $r>\operatorname{Max}((N+1)/2,n-\alpha(\mathcal{C}))$. Now $n-r<n-(n-\alpha(\mathcal{C}))=\alpha(\mathcal{C})$, hence $Z_{(1)}\,\varepsilon\,\mathcal{C},\cdots,Z_{(n-r)}\,\varepsilon\,\mathcal{C}$ by definition of $\alpha(\mathcal{C})$. Since $Z_{(s)}$ is finitely generated by 2.1 (a), we may use Theorem 2.3A:[3]

$$0\to H^{r+s}(K,L)\otimes Z_{(s)}\xrightarrow{\ \alpha\ }H^{r+s}(K,L;Z_{(s)})\xrightarrow{\ \beta\ }\operatorname{Tor}(H^{r+s+1}(K,L),Z_{(s)})\to 0$$

is an exact sequence. $Z_{(s)}\,\varepsilon\,\mathcal{C}$ for $s\leqq n-r$, hence $H^{r+s}(K,L;Z_{(s)})\,\varepsilon\,\mathcal{C}$ for $s\leqq n-r$ by properties (I) and ($\mathrm{II_B}$) of classes. For $s>n-r$,

$$H^{r+s}(K,L)\,\varepsilon\,\mathcal{C} \text{ and } H^{r+s+1}(K,L)\,\varepsilon\,\mathcal{C}$$

by hypothesis, and again by the above exact sequence $H^{r+s}(K,L;Z_{(s)})\,\varepsilon\,\mathcal{C}$ for $s>n-r$. This completes the proof.

---

[2] It is suggested that the reader draw a diagram of the relevant portion of the first derived cohomotopy exact couple to facilitate following the proof.

[3] 2.3A stands for Theorem 2.3 of the appendix.

Proof of 3.2. This proof is very similar to that of 3.1. Again we need to show that $H^{r+1}(K,L;Z_{(1)})\;\varepsilon\;\mathcal{C},\cdots,H^{N}(K,L;Z_{(N-r)})\;\varepsilon\;\mathcal{C}$ for $r\geqq n$. We use the above exact sequence. Since $H^{r+s}(K,L)\;\varepsilon\;\mathcal{C}$ for $s\geqq 1$ and $Z_{(s)}\;\varepsilon\;\mathcal{C}$ for $s\leqq N-n<N-(N-\alpha(\mathcal{C}))=\alpha(\mathcal{C})$, by conditions ($\mathrm{II_A}$) and (I) on classes the result is proven.

6. Proof of Theorem 3.9. Corollary 3.7 gives us the range of values of r for which $\eta^{r}{}_{(p)}$ is an isomorphism. The first open problems arising are determining the kernel of $\eta^{n-2p+3}{}_{(p)}$ and the image of $\eta^{n-2p+2}{}_{(p)}$ from the cohomology structure of the pair (K,L). This and more is achieved by the exact sequence of 3.9. In order to prove 3.9, we first prove a result which introduces the Steenrod reduced p-th power operations into the first derived cohomotopy exact couple.

These results lead us to the next open problem; namely, determining the kernel of $\eta^{n-4p+5}{}_{(p)}$ from the cohomology structure of the pair (K,L). This requires a further study of secondary (and higher order) cohomology operations. Some results on the nature of these operations, with applications to the above, have been obtained by the author and will appear at a later date.

LEMMA 6.1. *In the first derived cohomotopy exact couple, $\Gamma^{r+1,r+s}$ is a torsion group for $s\geqq 1$. Furthermore, $j'_{(p)}\colon\Gamma^{r+1,r+s}{}_p\to\Gamma^{r+1,r+s-1}{}_p$ is an isomorphism for $2\leqq s<2p-2$.*

Proof. The following exact sequence is part of the first derived cohomotopy exact couple:

$$H^{r+s-1}(K,L;Z_{(s-1)})\xrightarrow{\Delta'}\Gamma^{r+1,r+s}\xrightarrow{j'}\Gamma^{r+1,r+s-1}\xrightarrow{i'}H^{r+s}(K,L;Z_{(s-1)}).$$

By 2.1 (a), $Z_{(s-1)}\;\varepsilon\;\mathcal{C}_T$ for $s\geqq 2$, hence $H^{r+s-1}(K,L;Z_{(s-1)})\;\varepsilon\;\mathcal{C}_T$ for $s\geqq 2$. Also, $\Gamma^{r+1,N}=0\;\varepsilon\;\mathcal{C}_T$, therefore by induction, $\Gamma^{r+1,r+s-1}\;\varepsilon\;\mathcal{C}_T$ for $s\geqq 2$. Similarily, $Z_{(s-1)}\;\varepsilon\;\mathcal{C}_p$ for $2\leqq s<2p-2$ by 2.1 (b), hence $H^{r+s-1}(K,L;Z_{(s-1)})\;\varepsilon\;\mathcal{C}_p$ and $H^{r+s}(K,L;Z_{(s-1)})\;\varepsilon\;\mathcal{C}_p$ for $2\leqq s<2p-2$. Therefore, $j'_{(p)}\colon\Gamma^{r+1,r+s}{}_p\to\Gamma^{r+1,r+s-1}{}_p$ is an isomorphism for $2\leqq s<2p-2$. This completes the proof.

Let $r>(N+1)/2$. We define

$$d\colon H^{r}(K,L)\to H^{r+2p-2}(K,L;Z_{(2p-3)})_p=H^{r+2p-2}(K,L;Z_p)$$

by

$$d(u)=i'^{r+1,r+2p-2}{}_{(p)}(j'^{r+1,r+2p-3}{}_{(p)})^{-1}\cdots(j'^{r+1,r+2}{}_{(p)})^{-1}P\Delta'^{r,r}(u),$$

where $P\colon\Gamma^{r+1,r+1}\to\Gamma^{r+1,r+1}{}_p$ is the natural projection onto the p-primary component (P is naturally defined because $\Gamma^{r+1,r+1}$ is a torsion group).

THEOREM 6.2. *In the first derived cohomotopy exact couple, the homomorphism* $d\colon H^r(K,L)\to H^{r+2p-2}(K,L;Z_p)$ *is equal to* $\beta\mathcal{P}^1$, *where* $\beta\not\equiv 0$ $(\bmod p)$.

Proof. Since all the homomorphisms in the definition of d are natural with respect to cellular maps $f\colon (K,L)\to(K',L')$, d is also natural with respect to such maps and hence with respect to all maps (as in Section 4). Thus d is a cohomology operation and d corresponds to an element of $H^{r+2p-2}(Z,r;Z_p)\approx Z_p$ (see the calculations of Cartan [3]). Therefore $d=\beta\mathcal{P}^1$, and it suffices to exhibit a complex K for which $d\neq 0$ for then $\beta\not\equiv 0$ $(\bmod p)$.

Let $(K,L)=(M,x_0)$, where $M=S^r\cup e^{r+2p-2}$, the cell e^{r+2p-2} being attached to S^r by a non-zero element of $\pi_{r+2p-3}(S^r)_p=Z_p$. M has the property that $\mathcal{P}^1(u)\neq 0$, where u is a generator of $H^r(M,x_0)$ (this is a result of Borel and Serre [1; p. 425]). Assume $d=0$ for this complex. Since $H^s(M,x_0;G)=0$ unless $s=r$ or $s=r+2p-2$,

$$\Gamma^{r+1,r+1}\approx H^{r+2p-2}(M,x_0;Z_{(2p-3)}),$$

and $d(u)=P\Delta'^{r,r}(u)$ for $u\,\varepsilon\, H^r(M,x_0)$. Let u generate $H^r(M,x_0)\approx Z$. $\Delta'^{r,r}(u)$ is an element of finite order, hence there is an integer $D\equiv 1$ $(\bmod p)$ such that $P\Delta'^{r,r}(u)=D\Delta'^{r,r}(u)=\Delta'^{r,r}(Du)$. Since $d=0$, $\Delta'^{r,r}(Du)=0$, and by exactness, $Du=\eta^r([a])$, where $a\colon(M,x_0)\to(S^r,\text{pt.})$. $a\,|\,S^r\colon(S^r,x_0)\to(S^r,\text{pt.})$ is a map of degree D, and hence $a^*(u')=Du$, where u' is a generator of $H^r(S^r,\text{pt.})$. Thus

$$0\neq D\mathcal{P}^1(u)=\mathcal{P}^1(Du)=\mathcal{P}^1(a^*(u'))=a^*(\mathcal{P}^1(u'))=0$$

because $D\equiv 1$ $(\bmod p)$ and $\mathcal{P}^1(u')=0$ in S^r. This is a contradiction, and hence $d\neq 0$. This completes the proof.

Using a more computational proof, it can be shown that $\beta\equiv 1$ $(\bmod p)$. However, to prove Theorem 3.9, it is not necessary to know that $\beta\equiv 1$ $(\bmod p)$ because $\operatorname{Im}(\beta\mathcal{P}^1)=\operatorname{Im}\mathcal{P}^1$ and $\operatorname{Ker}(\beta\mathcal{P}^1)=\operatorname{Ker}\mathcal{P}^1$ as long as $\beta\not\equiv 0$ $(\bmod p)$. Using this remark, we now prove 3.9.

Proof of 3.9. The proof is based on the first derived cohomotopy exact couple.[2] By hypothesis, $r+4p-5>n$, and hence $H^{r+4p-5}(K,L;Z_{(4p-5)})\,\varepsilon\,\mathcal{C}_p$. It follows that $\Gamma^{r,r+2p-3}\,\varepsilon\,\mathcal{C}_p$ and $\Gamma^{r+1,r+2p-2}\,\varepsilon\,\mathcal{C}_p$. Hence

$$\Gamma^{r,r}{}_p\approx H^{r+2p-3}(K,L;Z_{(2p-3)})_p=H^{r+2p-3}(K,L;Z_p)$$

under the isomorphism $i'^{r,r+2p-3}{}_{(p)}(j'^{r,r+2p-4}{}_{(p)})^{-1}\cdots(j'^{r,r}{}_{(p)})^{-1}$. Furthermore,

$\operatorname{Im}(P\Delta'^{r-1,r-1}) = (\operatorname{Im}\Delta'^{r-1,r-1}) \cap \Gamma^{r,r}{}_p$. Similar remarks for $\Delta'^{r,r}$ and $\Gamma^{r+1,r+1}$ hold. From the exact couple, we obtain the exact sequence

$$H^{r-1}(K,L) \xrightarrow{P\Delta'^{r-1,r-1}} \Gamma^{r,r}{}_p \xrightarrow{j'^{r,r}{}_{(p)}} \pi^r(K,L)_p \xrightarrow{\eta^r{}_{(p)}} H^r(K,L)_p$$
$$\xrightarrow{\Delta'^{r,r}{}_{(p)}} \Gamma^{r+1,r+1}{}_p.$$

Using Theorem 6.2 and the above isomorphisms, we obtain the exact sequence of 3.9 with μ defined by

$$\mu = j'^{r,r}{}_{(p)} \cdot \cdot \cdot j'^{r,r+2p-4}{}_{(p)} (i'^{r,r+2p-3}{}_{(p)})^{-1}:$$
$$H^{r+2p-3}(K,L;Z_p) \to \pi^r(K,L)_p.$$

For $r = n - 4p + 5$, we have only the statement on the cokernel of $\eta^r{}_{(p)}$.

Appendix.

1. Results from Eilenberg and Steenrod [5]. The purpose of this appendix is to discuss the universal coefficient theorem for cohomology.

THEOREM 1.1. *Let K be a chain complex composed of free groups. For an arbitrary abelian group G, the following sequences are exact and split:*

$$0 \to H_r(K) \otimes G \xrightarrow{\alpha} H_r(K;G) \xrightarrow{\beta} \operatorname{Tor}(H_{r-1}(K), G) \to 0 \quad \textit{and}$$

$$0 \to \operatorname{Ext}(H_{r-1}(K), G) \xrightarrow{\beta} H^r(K;G) \xrightarrow{\alpha} \operatorname{Hom}(H_r(K), G) \to 0.$$

These exact sequences are natural with respect to chain maps $f\colon K \to K'$ and homomorphisms $\phi\colon G \to H$.

Proof. See exercise G-3 in [5; Chapt. V].

THEOREM 1.2. *Let K be a chain complex composed of finitely generated free groups. For an arbitrary abelian group G, the following sequences are exact and split:*

$$(*) \quad 0 \to H^r(K) \otimes G \xrightarrow{\alpha} H^r(K;G) \xrightarrow{\beta} \operatorname{Tor}(H^{r+1}(K), G) \to 0 \quad \textit{and}$$

$$0 \to \operatorname{Ext}(H^{r+1}(K), G) \xrightarrow{\beta} H_r(K;G) \xrightarrow{\alpha} \operatorname{Hom}(H^r(K), G) \to 0.$$

These exact sequences are natural with respect to chain maps $f\colon K \to K'$ and homomorphisms $\phi\colon G \to H$.

Proof. See exercises F-3, F-4, and G-3 in [5; Chapt. V].

Let $0 \to G' \xrightarrow{\phi} G \xrightarrow{\psi} G'' \to 0$ be an exact coefficient sequence.

$$\delta_*: H^r(K; G'') \to H^{r+1}(K; G')$$

is defined in [5; p. 158]. Define the sequence corresponding to this exact coefficient sequence to be the following sequence of groups and homomorphisms:

$$\cdots \to H^r(K; G') \xrightarrow{\phi_*} H^r(K; G) \xrightarrow{\psi_*} H^r(K; G'') \xrightarrow{\delta_*} H^{r+1}(K; G') \to \cdots.$$

THEOREM 1.3. *Let K be a chain complex composed of free groups. Then the sequence corresponding to the above exact coefficient sequence is exact. This exact sequence is natural with respect to chain maps $f: K \to K'$ and homomorphisms of one exact coefficient sequence into another.*

Proof. See exercise C-3 in [5; Chapt. V].

2. A new universal coefficient theorem. The sequence (*) of 1.2A does not necessarily hold if K is not finitely generated. We now prove a similar theorem by assuming G is finitely generated with K arbitrary.

LEMMA 2.1. *If G is finitely generated and free, then $\alpha: H^r(K) \otimes G \to H^r(K; G)$ is an isomorphism.*

Proof. α is obviously an isomorphism in case $G = Z$. Furthermore, the functors $H^r(K) \otimes G$ and $H^r(K; G)$ are additive with respect to G and thus commute with finite direct sums [4]. Hence α is an isomorphism if G is finitely generated and free.

THEOREM 2.2. *Let K be a chain complex composed of free groups. Let G be finitely generated. Then the sequence (*) of 1.2A is exact. This exact sequence is natural with respect to chain maps $f: K \to K'$ and homomorpisms $\phi: G \to H$.*

Proof. Let $0 \to R \xrightarrow{i} F \xrightarrow{j} G \to 0$ be exact, where F is a finitely generated free abelian group. R is finitely generated and free also. By 1.3A, the sequence corresponding to this exact coefficient sequence is exact:

$$\cdots \to H^r(K; R) \xrightarrow{i_*} H^r(K; F) \xrightarrow{j_*} H^r(K; G) \xrightarrow{\delta_*} H^{r+1}(K; R)$$

$$\xrightarrow{i_*} H^{r+1}(K; F) \to \cdots.$$

Hence the following sequence is exact:

$$0 \to \operatorname{Ker} \delta_* \to H^r(K; G) \to \operatorname{Im} \delta_* \to 0.$$

However, $\operatorname{Ker} \delta_* = \operatorname{Im} j_* \approx \operatorname{Coker} i_*$ and $\operatorname{Im} \delta_* = \operatorname{Ker} i_*$ by exactness. By 2.1A, we see that

$$\operatorname{Coker} i_* \approx \operatorname{Coker} (H^r(K) \otimes R \to H^r(K) \otimes F) \approx H^r(K) \otimes G$$

and

$$\operatorname{Ker} i_* \approx \operatorname{Ker} (H^{r+1}(K) \otimes R \to H^{r+1}(K) \otimes F) = \operatorname{Tor}(H^{r+1}(K), G)$$

(see [5; p. 160] for the definition of Tor). Under these isomorphisms, the inclusion $\operatorname{Ker} \delta_* \to H^r(K; G)$ goes over to α: $H^r(K) \otimes G \to H^r(K; G)$, and δ_* defines β: $H^r(K; G) \to \operatorname{Tor}(H^{r+1}(K), G)$. Hence the sequence (*) is exact. The naturality statements follow from the naturality statements of 1.3A.

COROLLARY 2.3. *The universal coefficient theorem for cohomology*

$$0 \to H^r(X, A) \otimes G \xrightarrow{\alpha} H^r(X, A; G) \xrightarrow{\beta} \operatorname{Tor}(H^{r+1}(X, A), G) \to 0$$

holds in the following cases:

1) *simplicial (or cellular) theory for finite complexes and G arbitrary, or not necessarily finite complexes and G finitely generated*:

2) *singular theory for G finitely generated*; *and*

3) *Cech theory for* (X, A) *compact and G arbitrary, or* (X, A) *paracompact and G fiinitely generated.*

Proof. This follows immediately from 1.2A and 2.2A and the fact that direct limits preserve exactness, $\otimes$, and Tor [4].

3. A counter-example. This example shows that the exact sequence of 2.3A does not hold in general for singular theory or cellular theory. Let X be an $(n-1)$-connected CW-complex such that $H_n(X) = Q =$ the additive group of rationals. Let $G = Q$. By 1.1A, $H^n(X; Z) = \operatorname{Hom}(Q, Z) = 0$, and $H^n(X; Q) = \operatorname{Hom}(Q, Q) = Q$. However, 2.3A would give that

$$0 \to 0 \otimes Q \to Q \to \operatorname{Tor}(H^{n+1}(X), Q) \to 0$$

is exact; but $\operatorname{Tor}(H^{n+1}(X), Q) = 0$, and hence $Q = 0$. This is a contradiction.

PRINCETON UNIVERSITY.

REFERENCES.

[1] A. Borel and J.-P. Serre, "Groupes de Lie et Puissances Réduites de Steenrod," *American Journal of Mathematics*, vol. 75 (1953), pp. 409-448.

[2] K. Borsuk, "Sur les groupes des classes de transformations continues," *Comptes Rendus de l'Academie Sciences*, Paris, vol. 202 (1936), pp. 1400-1403.

[3] H. Cartan, "Sur les groupes d'Eilenberg-MacLane," I and II, *Proceedings of the National Academy of Sciences*, vol. 40 (1954), pp. 467-471, 704-707.

[4] H. Cartan and S. Eilenberg, *Homological Algebra*, Princeton University Press, 1956.

[5] S. Eilenberg and N. E. Steenrod, *Foundations of Algebraic Topology*, Princeton University Press, 1953.

[6] P. J. Hilton, *An Introduction to Homotopy Theory*, Cambridge University Press, 1953.

[7] H. Hopf, "Die klassen der abbildungen der n-dimensionalen Polyeder der n-dimensionale Sphäre," *Commentarii Mathematici Helvetici*, vol. 5 (1953), pp. 39-54.

[8] W. S. Massey, "Exact couples in algebraic topology," I and II, III, IV, and V, *Annals of Mathematics*, vol. 56 (1952), pp. 363-396, vol. 57 (1953), pp. 248-286.

[9] W. S. Massey, "Some problems in algebraic topology and the theory of fibre bundles," *Annals of Mathematics*, vol. 62 (1955), pp. 327-359.

[10] J.-P. Serre, "Homologie singuliere des espaces fibrés," *Annals of Mathematics*, vol. 54 (1951), pp. 425-505.

[11] ———, "Groupes d'homotopie et classes de groupes Abéliens," *Annals of Mathematics*, vol. 58 (1953), pp. 258-294.

[12] ———, "Cohomologie modulo 2 des complexes d'Eilenberg-MacLane," *Commentarii Mathematici Helvetici*, vol. 27 (1953), pp. 198-323.

[13] E. H. Spanier, "Borsuk's cohomotopy groups," *Annals of Mathematics*, vol. 50 (1949), pp. 203-245.

[14] N. E. Steenrod, "Products of cocycles and extensions of mappings," *Annals of Mathematics*, vol. 48 (1947), pp. 290-320.

[15] ———, "Homology groups of symmetric groups and reduced power operations," and "Cyclic reduced powers of cohomology classes," *Proceedings of the National Academy of Sciences*, vol. 39 (1953), pp. 213-217, 217-223.

[16] G. W. Whitehead, "On the Freudenthal theorems," *Annals of Mathematics*, vol. 57 (1954), pp. 209-228.

GENERALIZED COHOMOTOPY GROUPS.*

By Franklin P. Peterson.[1]

1. Introduction. One of the fundamental problems of topology is the computation of $\pi(K;X)$, the set of homotopy classes of maps of a complex K into a space X. When K is an n-sphere S^n, then $\pi(K;X)=\pi_n(X)$, the familiar n-th homotopy group of X. When $X=S^n$ and the dimension of K is $\leqq 2n-2$, then $\pi(K;X)=\pi^n(K)$, the n-th cohomotopy group of K. The structure of $\pi^n(K)$ has been studied in [11].

In this paper, we shall introduce cohomotopy groups with coefficients in an abelian group G; namely, $\pi^n(K;G)=\pi(K;X)$, where the dimension of K is $\leqq 2n-2$ and X is a simply connected space whose homology is zero except that its n-th homology group is G. $\pi^n(K;G)$ is shown to be independent of the choice of such a space X. Notice that if $G=Z=$ the additive group of integers, then $\pi^n(K;Z)=\pi^n(K)$. For a fixed group G, the properties of $\pi^n(K;G)$ are analogous to those of $\pi^n(K)$, and the results of [11] are generalized so as to apply to $\pi^n(K;G)$. Furthermore, if G has no elements of order 2, then a homomorphism $\phi\colon G\rightarrow H$ induces a unique homomorphism $\phi_{\#}\colon \pi^n(K;G)\rightarrow\pi^n(K;H)$. Except for this restriction on G, $\phi_{\#}$ enjoys many of the same properties as the induced cohomology homomorphism $\phi_*\colon H^n(K;G)\rightarrow H^n(K;H)$. In particular, the sequence corresponding to an exact coefficient sequence is exact, and there is a universal coefficient theorem which asserts that

$$0\rightarrow\pi^n(K)\otimes G\xrightarrow{\alpha}\pi^n(K;G)\xrightarrow{\beta}\mathrm{Tor}(\pi^{n+1}(K),G)\rightarrow 0$$

is a split exact sequence. This theorem reduces the problem of computing $\pi^n(K;G)$ to that of computing $\pi^n(K)$. We conclude with a section on cohomotopy operations and with some remarks on homotopy groups with coefficients in G dual to our generalized cohomotopy groups.

In conclusion, I wish to express my warm appreciation to Professor N. E. Steenrod for his kind advice and encouragement. This paper is essentially Part II of a paper written under his direction and submitted as a disserta-

* Received September 21, 1955.

[1] The author was a predoctoral National Science Foundation Fellow during the preparation of this paper.

tion to Princeton University. I also wish to thank Professor J. C. Moore for suggesting the idea of general coefficients.

2. Homotopy classes of maps. In this section, we review some known results on the existence of a group structure on the set of homotopy classes of maps $a: (K, L) \to (X, x_0)$.

Let (K, L) be a CW-pair with dimension $K = N$ (K is an N-dimensional CW-complex [7] and L is a subcomplex). Let X be an arcwise connected space, and let $x_0 \varepsilon X$. We denote by $\pi(K, L; X, x_0)$ the set of homotopy classes of maps $a: (K, L) \to (X, x_0)$. A map $f: (K, L) \to (K', L')$ induces a function $f^{\#}: \pi(K', L'; X, x_0) \to \pi(K, L; X, x_0)$ defined by $f^{\#}([a]) = [af]$, where $[a]$ denotes the homotopy class of a. Also, a map $\phi: (X, x_0) \to X', x_0')$ induces a function $\phi_{\#}: \pi(K, L; X, x_0) \to \pi(K, L; X', x_0')$ defined by $\phi_{\#}([a]) = [\phi a]$.

If $X \neq \emptyset$, let SX denote the reduced suspension of X [15; p. 656]; namely, SX is the space obtained from $X \times I$ by identifying $X \times \{0\} \cup X \times \{1\} \cup \{x_0\} \times I$ to a point x_0. (x_0 is used to denote the base point of both X and SX.) Let $\mathcal{S}X$ denote the suspension of X [16]; namely, $\mathcal{S}X$ is the space obtained from $X \times I$ by identifying $X \times \{0\}$ and $X \times \{1\}$ to points. If $X = \emptyset$, define $S\emptyset$ and $\mathcal{S}\emptyset$ to be a pair of points. Also define $S^r X = S(S^{r-1}X)$, $S^0 X = X$, $\mathcal{S}^r X = \mathcal{S}(\mathcal{S}^{r-1}X)$, and $\mathcal{S}^0 X = X$. Let $\mu: (\mathcal{S}K, \mathcal{S}L) \to (SK, SL)$ be the canonical map contracting $\{x_0\} \times I$ to x_0. Since (K, L) is a CW-pair, μ is a homotopy equivalence and thus

$$\mu^{\#}: \pi(SK, SL; X, x_0) \to \pi(\mathcal{S}K, \mathcal{S}L; X, x_0)$$

is a 1-1 correspondence.

As in [16], suspension induces a function

$$S_{\#}: \pi(K, L; X, x_0) \to \pi(SK, SL; SX, x_0).$$

$S_{\#}$ is natural with respect to maps $f: (K, L) \to (K', L')$ and maps $\phi: (X, x_0) \to (X', x_0')$.

THEOREM 2.1. *If X is an $(n-1)$-connected space and (K, L) is a CW-pair with dimension $K \leqq 2n-2$, then $S_{\#}$ is a 1-1 correspondence.*

Proof. This is an immediate consequence of Corollary 7.2 of [16].

Let $(X, x_0)^{K,L}$ denote the function space of maps $a: (K, L) \to (X, x_0)$ with the compact-open topology. The constant map at x_0 serves as the base point for this function space. In [1], a function

$$\lambda: \pi_r((X, x_0)^{K,L}) \to \pi(S^r K, S^r L; X, x_0)$$

is defined and the following theorem is proven [1; p. 81]:

THEOREM 2.2. *λ is a 1-1 correspondence. This correspondence is natural with respect to maps $f\colon(K,L)\to(K',L')$ and maps $\phi\colon(X,x_0)\to(X',x_0')$.*

As in [16], we now prove the main result on the existence of a group structure on $\pi(K,L;X,x_0)$.

THEOREM 2.3. *If X is an $(n-1)$-connected space and (K,L) is a CW-pair with dimension $K\leqq 2n-2$, then $\pi(K,L;X,x_0)$ is an abelian group. This group structure is natural with respect to maps $f\colon(K,L)\to(K',L')$ and maps $\phi\colon(X,x_0)\to(X',x_0')$.*[2]

Proof. By 2.1, $S_{\#}{}^2\colon\ \pi(K,L;X,x_0)\to\pi(S^2K,S^2L;S^2X,x_0)$ is a natural 1-1 correspondence. Also, by 2.2,

$$\lambda\colon\ \pi_2((S^2X,x_0)^{K,L})\to\pi(S^2K,S^2L;S^2X,x_0)$$

is a natural 1-1 correspondence, and $\pi_2((S^2X,x_0)^{K,L})$ is an abelian group. We define the group structure on $\pi(K,L;X,x_0)$ using the 1-1 correspondence $\lambda^{-1}S_{\#}{}^2$. (This addition of homotopy classes is analogous to that defined in [14].) Note that $S_{\#}$ is now an isomorphism. This group structure is natural with respect to maps $f\colon(K,L)\to(K',L')$ and maps $\phi\colon(X,x_0)\to(X',x_0')$ because the 1-1 correspondences λ and $S_{\#}$ are; i.e. these maps induce homomorphisms $f^{\#}$ and $\phi_{\#}$.

We now define the sequence of a CW-pair (K,L). Define

$$\Delta\colon\ \pi(L;X,x_0)\to\pi(K,L;SX,x_0)$$

as follows: let $a\,\varepsilon\,[a]\,\varepsilon\,\pi(L;X,x_0)$. Extend a to a map $a'\colon(K,L)\to(CX,X)$, where CX denotes the cone on X [15; p. 656]. Let $h\colon(CX,X)\to(SX,x_0)$ be the canonical map collapsing X to x_0 [15; p. 657]. The composition $ha'\colon(K,L)\to(SX,x_0)$ represents $\Delta([a])$. When (K,L) and (X,x_0) satisfy the conditions of Theorem 2.3, Δ is a natural homomorphism (Δ is strictly analogous to the homomorphism Δ for ordinary cohomotopy [14; p. 216]). Let $i\colon L\to K$ and $j\colon K\to(K,L)$ be inclusions. Then the cohomotopy sequence of the pair (K,L) is defined to be the following sequence of groups and homomorphisms:

$$\pi(K,L;X,x_0)\xrightarrow{j^{\#}}\pi(K;X,x_0)\xrightarrow{i^{\#}}\pi(L;X,x_0)\xrightarrow{\Delta}\pi(K,L;SX,x_0)$$
$$\xrightarrow{j^{\#}}\pi(K;SX,x_0)\to\cdot\cdot\cdot.$$

[2] When we say a structure is natural with respect to maps $f\colon(K,L)\to(K',L')$ and maps $\phi\colon(X,x_0)\to(X',x_0')$, we assume (K',L') and (X',x_0') satisfy the same dimensional or connectedness assumptions that (K,L) and (X,x_0) satisfy.

In Section 4 we prove that this sequence is exact when (K,L) and (X,x_0) satisfy the conditions of Theorem 2.3.

3. Cohomotopy groups with coefficients in G. In this section we define our generalized cohomotopy groups and state their elementary properties. In making the definition certain arbitrary choices are necessary, and we prove independence of these choices under certain restrictions on G. We also state our main results on generalized cohomotopy groups; the proofs are given later in the paper.

Let G be an abelian group and let $n>1$. An $(n+1)$-dimensional CW-complex X is said to be an *$X(G,n)$-space* if $\pi_1(X)=0$, $H_i(X)=0$ for $i\neq n$, and $H_n(X)=G$. (This concept was introduced by Moore [10; p. 550].) Note that if X is an $X(G,n)$-space, then SX is an $X(G,n+1)$-space.

In Section 5 we prove

LEMMA 3.1. *For given G and n, there exists an $X(G,n)$-space.*

Let X be an $X(G,n)$-space, and let (K,L) be a CW-pair with dimension $K=N$. We define the *n-th cohomotopy group of (K,L) with coefficients in G* to be $\pi(K,L;X,x_0)$, and denote it by $\pi^n(K,L;G)$. When we use the notation $\pi^n(K,L;G)$, we assume $n>(N+1)/2$ and thus $\pi^n(K,L;G)$ has a natural group structure. As defined, $\pi^n(K,L;G)$ depends on the choice of $X(G,n)$-space. We show below that $\pi^n(K,L;G)$ is naturally independent of this choice when G has no elements of order 2. However, if we do not change coefficients during a discussion, it suffices to choose a fixed $X(G,t)$-space Y and use $S^{n-t}Y$ as the $X(G,n)$-space for each $n\geqq t\geqq 2$.

In Section 5 we prove

THEOREM 3.2. *$\pi^n(K,L;G)$ satisfies all the axioms for cohomology of Eilenberg and Steenrod* [6; p. 13] *in those dimensions where a natural group structure is defined.*

We now return to the question of independence of the choice of the $X(G,n)$-space. Let X be an $X(G,n)$-space, Y an $X(H,n)$-space. There is a natural homomorphism $\eta\colon \pi(X,x_0;Y,y_0)\to \mathrm{Hom}(G,H)$ defined by $\eta([a])=a_*\colon H_n(X,x_0)=G\to H_n(Y,y_0)=H$. In Section 5 we prove

THEOREM 3.3. *If $n\geqq 3$, then $\eta\colon \pi(X,x_0;Y,y_0)\to \mathrm{Hom}(G,H)$ is an epimorphism and has a kernel isomorphic to* $\mathrm{Ext}(G,H\otimes Z_2)$.[3]

Let $\mathfrak{D}$ be the family of abelian groups having no elements of order 2. In the appendix we prove

[3] See [5] or [6] for the definition and properties of Hom and Ext.

LEMMA 3.4. *If* $G \varepsilon \mathfrak{D}$, *then* $\mathrm{Ext}(G, H \otimes Z_2) = 0$.

Let $\phi: G \to H$ be a homomorphism, and let X be an $X(G,n)$-space, Y an $X(H,n)$-space. There exists a map $^*\phi: (X, x_0) \to (Y, y_0)$ such that $(^*\phi)_* = \phi: H_n(X, x_0) \to H_n(Y, y_0)$ by 3.3. In fact, if $G \varepsilon \mathfrak{D}$, then $^*\phi$ is unique up to homotopy by 3.3 and 3.4. Furthermore, if $\psi: H \to J$, Z is an $X(J,n)$-space, and $H \varepsilon \mathfrak{D}$ also, then $^*\psi^*\phi \simeq {}^*(\psi\phi): (X, x_0) \to (Z, z_0)$.

Now let X and X' be two different $X(G,n)$-spaces, let $G \varepsilon \mathfrak{D}$, and let $\phi: G \to G$ be the identity homomorphism. Then there exist maps $^*\phi: (X, x_0) \to (X', x_0')$ and $^*\phi': (X', x_0') \to (X, x_0)$ inducing ϕ such that $(^*\phi')(^*\phi)$ and $(^*\phi)(^*\phi')$ are homotopic to the identity maps. Furthermore, $^*\phi$ and $^*\phi'$ are unique up to homotopy. Hence $^*\phi$ and $^*\phi'$ induce unique isomorphisms

$$(^*\phi)_\#: \pi(K, L; X, x_0) \to \pi(K, L; X', x_0')$$

and

$$(^*\phi')_\#: \pi(K, L; X', x_0') \to \pi(K, L; X, x_0)$$

which are inverses of each other. Hence the set of groups $\{\pi(K, L; X, x_0)\}$ for all $X(G,n)$-spaces X form a transitive system of groups [6; p. 17], and we have shown that $\pi^n(K, L; G)$ is independent of the choice of $X(G,n)$-space. In case $G \not\varepsilon \mathfrak{D}$, $(^*\phi)_\#$ is an isomorphism, but it is not a unique isomorphism. In this case, we assume a fixed $X(G,2)$-space during any given discussion.

As a further corollary of the above discussion, if $G \varepsilon \mathfrak{D}$, then a homomorphism $\phi: G \to H$ induces a unique homomorphism $\phi_\#: \pi^n(K, L; G) \to \pi^n(K, L; H)$. This is natural in the sense that if $\phi: G \to H$, $\psi: H \to J$ and G and $H \varepsilon \mathfrak{D}$, then $(\psi\phi)_\# = \psi_\#\phi_\#$.

We now state our main results on cohomotopy groups with coefficients in G. Let G' and $G \varepsilon \mathfrak{D}$, and let $0 \to G' \xrightarrow{\phi} G \xrightarrow{\psi} G'' \to 0$ be an exact coefficient sequence. In Section 6 we define a homomorphism

$$\delta_\#: \pi^r(K, L; G'') \to \pi^{r+1}(K, L; G')$$

for $r > (N+1)/2$. Define the sequence corresponding to this exact coefficient sequence to be the following sequence of groups and homomorphisms:

$$\cdots \to \pi^r(K, L; G') \xrightarrow{\phi_\#} \pi^r(K, L; G) \xrightarrow{\psi_\#} \pi^r(K, L; G'')$$
$$\xrightarrow{\delta_\#} \pi^{r+1}(K, L; G') \to \cdots.$$

In Section 6 we prove

THEOREM 3.5. *For $r > (N+1)/2$, the sequence corresponding to an exact coefficient sequence is exact. This sequence is natural with respect to maps $f: (K,L) \to (K',L')$, and if G', G, G'', H', and $H \varepsilon \mathfrak{D}$, then it is natural with respect to a homomorphism of one exact sequence into another*:

$$\begin{array}{ccccccccc} 0 \to & G' & \xrightarrow{\phi} & G & \xrightarrow{\psi} & G'' & \to 0 \\ & \downarrow & & \downarrow & & \downarrow & \\ 0 \to & H' & \xrightarrow{\phi'} & H & \xrightarrow{\psi'} & H'' & \to 0. \end{array}$$

The above is a generalization of an exact sequence of Moore [10; p. 552].

Define a function $\pi^r(K,L) \times G \to \pi^r(K,L;G)$ by $([a],[g]) \to [ga]$, where $[g] \varepsilon \pi_r(X,x_0) = G$, $[a] \varepsilon \pi^r(K,L)$, and X is an $X(G,r)$-space. By 2.3, this is a bilinear function for $r > (N+1)/2$, and hence it induces a homomorphism α: $\pi^r(K,L) \otimes G \to \pi^r(K,L;G)$. In Section 7 we define a homomorphism β: $\pi^r(K,L;G) \to \operatorname{Tor}(\pi^{r+1}(K,L),G)$ and prove the following universal coefficient theorems:

THEOREM 3.6. *Let G be a finitely generated abelian group. Then the sequence*

$$(*) \qquad 0 \to \pi^r(K,L) \otimes G \xrightarrow{\alpha} \pi^r(K,L;G) \xrightarrow{\beta} \operatorname{Tor}(\pi^{r+1}(K,L),G) \to 0$$

is exact for $r > (N+1)/2$. This sequence is natural with respect to maps f: $(K,L) \to (K',L')$, and if $G \varepsilon \mathfrak{D}$, then it is natural with respect to homomorphisms $\phi: G \to H$. Furthermore, if $\pi^r(K,L)$ is finitely generated and $G \varepsilon \mathfrak{D}$, then the exact sequence (*) splits.

THEOREM 3.7. *Let $G \varepsilon \mathfrak{D}$, and let (K,L) be a finite CW-pair. Then the sequence (*) is exact for $r > (N+1)/2$. This sequence is natural with respect to maps $f: (K,L) \to (K',L')$ and homomorphisms ϕ: $G \to H$.*

In Section 7 we give an example to show that if G is not finitely generated and (K,L) is not finite, then the sequence (*) is not necessarily exact.

Theorem 3.6 generalizes an exact sequence of Serre [13; p. 284]; Serre's sequence is 3.6 for the case $(K,L) = (S^n, \text{pt.})$ and $G =$ a cyclic group. He notes that his sequence does not split when $G = Z_2$. It is a result of Barratt [2; p. 283] that Serre's sequence splits for G a cyclic group other than Z_2.

Theorems 3.6 and 3.7, beside giving a further analogy between cohomotopy groups with arbitrary coefficients and cohomology groups with arbitrary coefficients, reduce the problem of calculating $\pi^r(K,L;G)$ to the standard

problem of calculating $\pi^r(K,L)$. Hence, Theorems 3.6 and 3.7 are two of our main results.

Let X be an $X(G,n)$-space. $S_{\#}: \pi_{n+s}(X) \to \pi_{n+s+1}(SX)$ is an isomorphism when $s < n-1$ by Theorem 2.1. We identify these groups under this isomorphism and denote the result by $G_{(s)}$. For any class $\mathcal{C}$ [13], let $\alpha(\mathcal{C};G)$ denote the largest integer such that $G_{(s)} \varepsilon \mathcal{C}$ for $0 < s < \alpha(\mathcal{C};G)$ (see definition in [11]). In Section 4 we define a natural homomorphism $\eta^r: \pi^r(K,L;G) \to H^r(K,L;G)$ analogous to η^r for ordinary cohomotopy. In Section 8 we prove the following generalization of Theorem 3.1 of [11]:

THEOREM 3.8. *Let $\mathcal{C}$ be a class satisfying condition* ($\mathrm{II_B}$) *(see* [13] *or* [11]), *and let G be finitely generated. Let $n > (N+1)/2$ be such that $H^r(K,L;G) \varepsilon \mathcal{C}$ for every $r > n$. Then $\eta^r: \pi^r(K,L;G) \to H^r(K,L;G)$ is a $\mathcal{C}$-isomorphism if $r > \mathrm{Max}((N+1)/2, n-\alpha(\mathcal{C};G))$, and is a $\mathcal{C}$-epimorphism for $r = n-\alpha(\mathcal{C};G)$ in case $n-\alpha(\mathcal{C};G) > (N+1)/2$.*

Theorems 2.1 of [11] and 3.6 give information on $\alpha(\mathcal{C};G)$ for various $\mathcal{C}$ and G, and we may draw consequences of 3.8 similar to 3.2, 3.3, 3.4, and 3.5 of [11]. Also, the result of Adem carries over to general coefficients and there is a theorem analogous to Theorem 3.8 of [11]. Furthermore, for the case $G = Z_p$, there is a theorem analogous to Theorem 3.9 of [11]. Rather than considering these in detail, let us note that any result on the structure of $\pi^r(K,L)$ gives a result on the structure of $\pi^r(K,L;G)$ by 3.6 or 3.7.

4. The exact sequences of a pair. In this section we prove the exactness of the cohomotopy sequence of a pair defined in Section 2. We then show that the cohomotopy exact couple of [11] can be generalized; some of our main results are based on this generalized cohomotopy exact couple.

In the cohomotopy sequence of the pair (K,L), the following relations are obvious: $\mathrm{Im}\, j^{\#} = \mathrm{Ker}\, i^{\#}$, $\mathrm{Im}\, i^{\#} \subset \mathrm{Ker}\, \Delta$, and $\mathrm{Im}\, \Delta \subset \mathrm{Ker}\, j^{\#}$. To prove $\mathrm{Im}\, i^{\#} = \mathrm{Ker}\, \Delta$ and $\mathrm{Im}\, \Delta = \mathrm{Ker}\, j^{\#}$, we need the assumptions that X is $(n-1)$-connected and dimension $K = N \leqq 2n-2$. The exactness now follows from Theorem $(3.1)_0$ of [15; p. 657] with the carrier ϕ from K to $S^{N-n+1}X$ being such that $\phi L = x_0$ and with the unrestricted carrier ψ from K to $S^{N-n+1}X$. The complete details are left to the reader. Note that the exact sequence of a pair (K,L) gives immediately the exact cohomotopy sequence of a triple (K,L,M) [6; p. 25].

The results of Section 4 of [11] on the cohomotopy exact couple now carry over to our more general situation. Let X be an $(n-1)$-connected space. We replace S^{n+t} in [11] by S^tX for $t \geqq 0$. With this substitution,

all of the results of Section 4 of [11] are true; the main identifications now being that $\mathcal{H}^{r,s} \approx H^s(K,L;\pi_{s-r+n}(X))$ for $r > \mathrm{Max}(n, (N+1)/2)$ and $\Gamma^{r,r-1} \approx \pi(K,L;S^{r-n}X,x_0)$ for $r > \mathrm{Max}(n-1, (N+1)/2)$. Again we denote $i'^{r,r}$: $\pi(K,L;S^{r-n}X,x_0) \to H^r(K,L;\pi_n(X))$ by η^r.

For later use, we make the following definition. Let X be an $(n-1)$-connected space. Let dimension $K = N \leqq 2(s-r+n)-2$ and let $s \geqq r$. We define

$$\theta\colon \pi_r((S^sX,x_0)^{K,L}) \to \pi(K,L;S^{s-r}X,x_0)$$

to be the following composition:

$$\pi_r((S^sX,x_0)^{K,L}) \xrightarrow{\lambda} \pi(\mathcal{S}^rK,\mathcal{S}^rL;S^sX,x_0) \xrightarrow{\mu^{\#-1}} \pi(S^rK,S^rL;S^sX,x_0)$$
$$\xrightarrow{(S_\#^{-1})^r} \pi(K,L;S^{s-r}X,x_0).$$

λ is the isomorphism of 2.2, $S_\#^{-1}$ is the isomorphism of 2.1 ($S_\#$ is an isomorphism under the above restrictions on s, r, n, and N), and $\mu^\#$ is the isomorphism of Section 2. θ is natural with respect to maps $f\colon (K,L) \to (K',L')$ and maps $\phi\colon (X,x_0) \to (X',x_0')$ because λ, $\mu^\#$, and $S_\#$ are.

5. Independence of the choice of $X(G,n)$-space. In this section we give proofs of 3.1, 3.2, and 3.3 as well as giving some further properties of $\pi^n(K,L;G)$. These further properties will be used in Section 7.

Proof of 3.1. Let $G = F/R$, where F is a free abelian group on generators $\{x_\alpha\}_{\alpha\varepsilon A}$ and R is the group of relations. R is free abelian [6; p. 134] with basis $\{y_\beta\}_{\beta\varepsilon B}$. Let T be the CW-complex $\bigvee_{\alpha\varepsilon A} S^n{}_\alpha$, a union of n-spheres $S^n{}_\alpha$ with a single point in common. By the Hurewicz theorem, $\pi_n(T) \overset{\eta}{\approx} H_n(T) = F$. For each $\beta\varepsilon B$, attach an $(n+1)$-cell e_β to T by a map representing $\eta^{-1}(y_\beta)\,\varepsilon\,\pi_n(T)$. Call the resulting space X. By construction, $\pi_1(X) = 0$, $H_i(X) = 0$ for $i < n$, $\pi_n(X) \approx H_n(X) = G$, and $H_i(X) = 0$ for $i > n+1$. Moreover, since any non-zero $(n+1)$-cycle in X would imply a non-trivial relation among the $\{y_\beta\}_{\beta\varepsilon B}$, $H_{n+1}(X) = 0$.

In order to prove 3.3, we first prove

LEMMA 5.1. *Let X be an $X(G,n)$-space. Then $\pi_{n+1}(X) \approx G \otimes Z_2$.*

Proof. The proof is based on figure 1, a portion of the first derived homotopy exact couple of Massey [8; part II]. $H_{n+1}(X) = 0$ and hence by

exactness $\pi_{n+1}(X, x_0) \approx \Gamma \approx H_n(X; Z_2)$. By this universal coefficient theorem, $H_n(X; Z_2) \approx G \otimes Z_2$.

$$
\begin{tikzcd}
 & & \vdots \arrow[d] & & & \\
 & \cdots \arrow[r] & 0 \arrow[r] \arrow[d] & \cdots & \vdots \arrow[d] & \\
\cdots \arrow[r] & 0 \arrow[r] & \Gamma \arrow[r] \arrow[d] & H_n(X; Z_2) \arrow[r] & 0 \arrow[r] \arrow[d] & \cdots \\
 & \cdots \arrow[r] & \pi_{n+1}(X, x_0) \arrow[r] \arrow[d] & H_{n+1}(X) \arrow[r] & 0 \arrow[r] \arrow[d] & \cdots \\
 & & \vdots & & \vdots &
\end{tikzcd}
$$

Figure 1.

Proof of 3.3. The proof is based on figure 2, a portion of the first derived generalized cohomotopy exact couple. Let $(K, L) = (X, x_0)$, $N = n + 1$, and let (Y, y_0) be the coefficient space. Now

$$H^n(X, x_0; \pi_n(Y, y_0)) \approx H^n(X, x_0; H) \approx \operatorname{Hom}(G, H)$$

by the universal coefficient theorem. By 5.1 and the universal coefficient theorem,

$$H^{n+1}(X, x_0; \pi_{n+1}(Y, y_0)) \approx H^{n+1}(X, x_0; H \otimes Z_2) \approx \operatorname{Ext}(G, H \otimes Z_2).$$

By exactness η is an epimorphism with kernel isomorphic to

$$\Gamma^{n,n} \approx \operatorname{Ext}(G, H \otimes Z_2).$$

$$
\begin{tikzcd}
 & & \vdots \arrow[d] & & & \\
 & \cdots \arrow[r] & 0 \arrow[r] \arrow[d] & \cdots & \vdots \arrow[d] & \\
\cdots \arrow[r] & 0 \arrow[r] & \Gamma^{n,n} \arrow[r, "i'"] \arrow[d, "j'"] & H^{n+1}(X, x_0; \pi_{n+1}(Y, y_0)) \arrow[r] & 0 \arrow[r] \arrow[d] & \cdots \\
 & \cdots \arrow[r] & \pi(X, x_0; Y, y_0) \arrow[r, "\eta"] \arrow[d] & H^n(X, x_0; \pi_n(Y, y_0)) \arrow[r] & 0 \arrow[r] \arrow[d] & \cdots \\
 & & \vdots & & \vdots &
\end{tikzcd}
$$

Figure 2.

Before we prove 3.2, let us draw some further corollaries of 3.3.

COROLLARY 5.2. *If $G \varepsilon \mathfrak{D}$, then $\pi^n(K, L; G)$ is a unitary left module over any subring of the ring of endomorphisms of G.*

Proof. By the naturality statements of Theorem 2.3, $\pi^n(K,L;G)$ is a unitary left module over $\pi(X,x_0;X,x_0)$ where X is an $X(G,n)$-space (the multiplication is composition). By 3.3 and 3.4, $\eta\colon \pi(X,x_0;X,x_0) \to \mathrm{Hom}(G,G)$ is an isomorphism if $G \varepsilon \mathfrak{D}$, and the corollary follows.

COROLLARY 5.3. *If G is a field and $G \varepsilon \mathfrak{D}$, then $\pi^n(K,L;G)$ is a vector space over G.*

Proof. Any field is a subring of its ring of additive endomorphisms; namely, $g \varepsilon G$ acts on G by left multiplication. The corollary now follows from Corollary 5.2.

We now complete the proof of 3.2. The results of Sections 2 and 4 show that $\pi^n(K,L;G)$ satisfy axioms 1 through 4 of Eilenberg and Steenrod.

Let (K,L) be a CW-pair with $L \neq \emptyset$. Let (K_L, p_L) be the CW-pair obtained by identifying L to a point p_L. Let $f\colon (K,L) \to (K_L, p_L)$ be the canonical map.

LEMMA 5.4. *$f^{\#}\colon \pi(K_L,p_L;X,x_0) \to \pi(K,L;X,x_0)$ is a 1-1 correspondence.*

Proof. Same as in [14; p. 215].

Let (K,L) be a CW-pair. Let $M \subset L$ be such that $K - M$ is a subcomplex of K. Let $i\colon (K-M, L-M) \to (K,L)$ be the inclusion.

THEOREM 5.5. (Excision Axiom).

$$i^{\#}\colon \pi^r(K,L;G) \to \pi^r(K-M, L-M;G)$$

is an isomorphism for $r > (N+1)/2$.

Proof. This follows immediately from 5.4 as in [14; p. 215].

Let $f, g\colon (K,L) \to (K',L')$ be homotopic.

THEOREM 5.6. (Homotopy Axiom).

$$f^{\#} \text{ and } g^{\#}\colon \pi^r(K',L';G) \to \pi^r(K,L;G)$$

are equal.

Proof. Obvious.

Theorem 3.2 has now been proven. A corollary to this theorem is the fact that any theorem derivable from the axioms for cohomology of Eilenberg and Steenrod which does not make use of the lower dimensional groups holds for cohomotopy groups with coefficients in G. An example of this is the Mayer-Vietoris sequence of a triad [6; p. 39].

6. Proof of Theorem 3.5. This section is devoted to the proof that the sequence corresponding to an exact coefficient sequence is exact. The preliminary results are only steps in this proof.

LEMMA 6.1. *Let $G \varepsilon \mathfrak{D}$, and let $n > (N+1)/2$. Let X be an $X(G,n)$-space. Let Y be an $(n-1)$-connected space such that $H_n(Y, y_0) = G$ and $H_r(Y, y_0) = 0$ for $n < r < 2n$. Then there is a map $k: (X, x_0) \to (Y, y_0)$, unique up to homotopy, such that $k_*: H_n(X, x_0) = G \to H_n(Y, y_0) = G$ is the identity and $k_\#: \pi^n(K, L; G) \to \pi(K, L; Y, y_0)$ is a natural isomorphism.*

Proof. Similarly to the argument used in the proof of 3.3, there is a map $k: (X, x_0) \to (Y, y_0)$ such that $k_*: H_n(X, x_0) = G \to H_n(Y, y_0) = G$ is the identity. Hence $k_*: H_r(X, x_0) \to H_r(Y, y_0)$ is an isomorphism for $r < 2n$; it follows by a theorem of J. H. C. Whitehead [18; p. 215] that $k_\#: \pi_r(X, x_0) \to \pi_r(Y, y_0)$ is an isomorphism for $r < 2n-1$ and an epimorphism for $r = 2n-1$. Assuming that k is an inclusion by the mapping cylinder construction, this means that the pair (Y, X) is $(2n-1)$-connected [3; p. 183]. The lemma is now as easy consequence of the deformation obstruction theory as described in [3; p. 186]. The details of the proof are left to the reader.

Let $p: E \to B$ be a fibre space in the sense of Serre, with fibre F over $b_0 \varepsilon B$. Let K be a CW-complex. Define a map $p_1: E^K \to B^K$ by $p_1(a) = pa$.

LEMMA 6.2. *$p_1: E^K \to B^K$ is a fibre space with fibre F^K. A map $f: K \to K'$ induces a fibre preserving map $^*f: (E^{K'}, p_1', B^{K'}) \to (E^K, p_1, B^K)$, and a fibre preserving map $\phi: (E, p, B) \to (E', p', B')$ induces a fibre preserving map $^*\phi: (E^K, p_1, B^K) \to (E'^K, p_1', B'^K)$.*

Proof. The proof is a straight forward application of the covering homotopy theorem for (E, p, B), and the details are omitted.

Let G' and $G \varepsilon \mathfrak{D}$, and let $0 \to G' \xrightarrow{\phi} G \xrightarrow{\psi} G'' \to 0$ be an exact coefficient sequence. Let X be an $X(G, N+1)$-space, and let X'' be an $X(G'', N+1)$-space. Let $^*\psi: (X, x_0) \to (X'', x_0'')$ be a map inducing ψ. We replace $^*\psi$ by a fibre mapping as follows: assume $^*\psi$ is an inclusion by the mapping cylinder construction. Let Y be the space of paths in X'' which end in X. X is contained in Y as a deformation retract [6; p. 30] by $x \to$ constant path at x. Define $p: Y \to X''$ by $p(f) = f(0)$; p is a fibre map [12; p. 479] with fibre $F' =$ the space of paths starting at $x_0 \varepsilon X''$ and ending in X. We may assume $x_0 \varepsilon X \subset X''$. p is our replacement of $^*\psi$.

Using the technique of spectral sequences as applied to fibre spaces by Serre, we now prove the following lemma.

LEMMA 6.3. *$H_{N+1}(F') = G'$, $H_r(F') = 0$ otherwise for $r < 2(N+1) - 1$.*

Proof. Let E be the space of paths in X'' starting at x_0; E is a contractible space. Let $p_1 : (E, F') \to (X'', X)$ be defined by $p_1(f) = f(1)$. This is a relative fibre space with fibre $F =$ the space of loops in X'' at x_0. F is $(N-1)$-connected because X'' is N-connected. As in [9; p. 330], there is a spectral sequence of this relative fibre space with $E_2^{p,q} \approx H_p(X'', X; H_q(F))$ and E_∞ is the graded group associated with $H(E, F')$. From the exact homology sequence of the pair (X'', X):

$$\cdots \to H_{N+2}(X'') \to H_{N+2}(X'', X) \to H_{N+1}(X) \to H_{N+1}(X'') \\ \to H_{N+1}(X'', X) \to H_N(X) \to \cdots$$

we see that $H_{N+2}(X'', X) = G'$ and $H_r(X'', X) = 0$ otherwise. Hence $E_2^{p,q} = 0$ for $q < N$ except that $E_2^{N+2,0} = G'$, and $E_2^{p,q} = 0$ for $p < N+2$. Thus $H_r(E, F') = 0$ for $r < 2N + 2$ except that $H_{N+2}(E, F') = G'$. However E is contractible, and hence by the exact homology sequence of the pair (E, F'), $H_{N+1}(F') = G'$ and $H_r(F') = 0$ otherwise for $r < 2N + 1$.

In order to prove the naturality of Theorem 3.5, we first prove the following lemma.

LEMMA 6.4. *Given a homomorphism of one exact coefficient sequence into another:*

$$\begin{array}{ccccccccc} 0 \to & G' & \xrightarrow{\phi} & G & \xrightarrow{\psi} & G'' & \to 0 \\ & \downarrow{\scriptstyle \xi'} & & \downarrow{\scriptstyle \xi} & & \downarrow{\scriptstyle \xi''} & \\ 0 \to & H' & \xrightarrow{\phi'} & H & \xrightarrow{\psi'} & H'' & \to 0. \end{array}$$

Let G', G, G'', H', and $H \in \mathfrak{D}$. Refer to the above construction. Then there is a fibre preserving map $\xi : (Y, p, X'') \to (Y_H, p, X''_H)$ which is homotopic to $*\xi$ on Y and to $*\xi''$ on X''.

Proof.

$$\begin{array}{ccc} G & \xrightarrow{\psi} & G'' \\ \downarrow{\scriptstyle \xi} & & \downarrow{\scriptstyle \xi''} \\ H & \xrightarrow{\psi'} & H'' \end{array} \quad \text{is commutative and hence} \quad \begin{array}{ccc} X & \xrightarrow{*\psi} & X'' \\ \downarrow{\scriptstyle *\xi} & & \downarrow{\scriptstyle *\xi''} \\ X_H & \xrightarrow{*\psi'} & X''_H \end{array}$$

is commutative up to homotopy by the results of Section 3. Assume $*\psi$ and $*\psi'$ are inclusions. $*\xi : X \to X_H$ is such that $(*\psi')(*\xi) \simeq (*\xi'')(*\psi)$, i.e. $*\xi$ is homotopic in X''_H to a map which can be extended to all of X'', namely $*\xi'' | X$. Use the homotopy extension theorem to define a map $\xi_1 : (X'', X)$

$\to (X''_H, X_H)$ such that $\xi_1|X'' \simeq *\xi''$ and $\xi_1|X \simeq *\xi$. ξ_1 induces a fibre preserving map $\xi\colon (Y, p, X'') \to (Y_H, p, X''_H)$ having the correct properties. This completes the proof.

We now prove Theorem 3.5.

Proof of 3.5. By 5.4 and the exact cohomotopy sequence of the pair (K_L, p_L), it suffices to prove this theorem for the case $L = \emptyset$. Use the above construction and 6.2 to obtain the fibre space (Y^K, p_1, X''^K) with fibre F'^K. Consider the homotopy sequence of this fibre space and use the isomorphism θ of Section 4:

$$\begin{array}{ccccccccc}
\cdots \to \pi_{N+1-r}(F'^K) & \xrightarrow{j'_\#} & \pi_{N+1-r}(Y^K) & \xrightarrow{p_{1\#}} & \pi_{N+1-r}(X''^K) & \xrightarrow{\partial} & \pi_{N-r}(F'^K) \to \cdots \\
\downarrow k_\#^{-1}\theta & & \downarrow \theta & & \downarrow \theta & & \downarrow k_\#^{-1}\theta \\
\cdots \to \pi^r(K;G') & \xrightarrow{\phi_\#} & \pi^r(K;G) & \xrightarrow{\psi_\#} & \pi^r(K;G) & \xrightarrow{\delta_\#} & \pi^{r+1}(K;G') \to \cdots
\end{array}$$

∂ is the boundary operator in the homotopy sequence of the fibre space, $j'_\#$ is induced by the inclusion $j'\colon F'^K \to Y^K$, and $p_{1\#}$ is induced by p_1. θ is defined and is an isomorphism for $r > (N+1)/2$ because

$$2(N+1-N-1+r)-2 = 2r-2 > 2(N+1)/2-2 = N-1.$$

$k_\#\colon \pi^r(K;G') \to \pi(K;F')$ is the isomorphism of 6.1 and 6.3, and $\delta_\#\colon \pi^r(K;G'') \to \pi^{r+1}(K;G')$ is defined by $\delta_\# = k_\#^{-1}\theta\partial\theta^{-1}$. Furthermore, $\theta j'_\# = \phi_\# k_\#^{-1}\theta$ and $\psi_\#\theta = \theta p_{1\#}$ by the naturality of θ. Also, $\delta_\#\theta = k_\#^{-1}\theta\partial$ by definition. Hence the sequence corresponding to the exact coefficient sequence is exact.

A map $f\colon K \to K'$ induces a fibre preserving map

$$f\colon (Y^{K'}, p_1', X''^{K'}) \to (Y^K, p_1, X''^K)$$

and hence $f^\#$ commutes with $\delta_\#$ because ∂ and θ are natural. $f^\#$ commutes with $\phi_\#$ and $\psi_\#$ also. Thus f induces a homomorphism of the sequence of K' into that of K. Also, a homomorphism of one exact coefficient sequence into another

$$\begin{array}{ccccccccc}
0 \to & G' & \xrightarrow{\phi} & G & \xrightarrow{\psi} & G'' & \to 0 \\
& \downarrow \xi' & & \downarrow \xi & & \downarrow \xi'' & \\
0 \to & H' & \xrightarrow{\phi'} & H & \xrightarrow{\psi'} & H'' & \to 0,
\end{array}$$

induces a fibre preserving map $\xi\colon (Y^K, p_1, X''^K) \to (Y_H{}^K, p_1, X''_H{}^K)$ by Lemma 6.4 when G', G, G'', H', and $H \in \mathfrak{D}$. Hence $\xi_\#$ commutes with $\delta_\#$ as well as

with $\phi_\#$ and $\psi_\#$. Thus a homomorphism of the exact coefficient sequence induces a homomorphism of the sequence corresponding to that exact coefficient sequence. This completes the proof.[4]

7. Proof of the universal coefficient theorem. This section is devoted to proving that the sequence

$$(*)\qquad 0 \to \pi^r(K,L)\otimes G \xrightarrow{\alpha} \pi^r(K,L;G) \xrightarrow{\beta} \operatorname{Tor}(\pi^{r+1}(K,L),G) \to 0$$

is exact under various hypotheses.

LEMMA 7.1. *$\alpha\colon \pi^r(K,L)\otimes G \to \pi^r(K,L;G)$ is natural with respect to maps $f\colon (K,L)\to(K',L')$, and if $G\,\varepsilon\,\mathfrak{D}$, then it is natural with respect to homomorphisms $\phi\colon G\to H$. Moreover, if G is finitely generated and free, then α is an isomorphism.*

Proof. The naturality statements are obvious. If $G=Z$, α is obviously an isomorphism because $\pi^r(K,L;Z)=\pi^r(K,L)$. Moreover, the functors $\pi^r(K,L)\otimes G$ and $\pi^r(K,L;G)$ are additive [5] on $\mathfrak{D}$ by the results of Section 3. Hence they commute with finite direct sums, and α is an isomorphism if G is finitely generated and free.

Proof of the first part of 3.6. Let $0\to R\xrightarrow{i} F\xrightarrow{j} G\to 0$ be exact, where F is a finitely generated free abelian group. R is finitely generated and free also. By 3.5, the sequence corresponding to this coefficient sequence is exact (R and $F\,\varepsilon\,\mathfrak{D}$):

$$\cdots\to\pi^r(K,L;R)\xrightarrow{i_\#}\pi^r(K,L;F)\xrightarrow{j_\#}\pi^r(K,L;G)\xrightarrow{\delta_\#}\pi^{r+1}(K,L;R)$$
$$\xrightarrow{i_\#}\pi^{r+1}(K,L;F)\to\cdots.$$

Hence the following sequence is exact:

$$0\to\operatorname{Ker}\delta_\#\to\pi^r(K,L;G)\to\operatorname{Im}\delta_\#\to 0.$$

However, $\operatorname{Ker}\delta_\#=\operatorname{Im}j_\#\approx\operatorname{Coker}i_\#$ and $\operatorname{Im}\delta_\#=\operatorname{Ker}i_\#$ by exactness. By 7.1, we see that

$$\operatorname{Coker} i_\#\approx\operatorname{Coker}(\pi^r(K,L)\otimes R\to\pi^r(K,L)\otimes F)\approx\pi^r(K,L)\otimes G$$

and

$$\operatorname{Ker} i_\#\approx\operatorname{Ker}(\pi^{r+1}(K,L)\otimes R\to\pi^{r+1}(K,L)\otimes F)=\operatorname{Tor}(\pi^{r+1}(K,L),G)$$

[4] The reader who is familiar with the techniques of Spanier and Whitehead can construct an alternate proof of Theorem 3.5 using those techniques.

(see [5] for the elementary properties of Tor needed in this proof). Under these isomorphisms, the inclusion $\operatorname{Ker}\delta_{\#} \to \pi^r(K,L;G)$ goes over to $\alpha\colon \pi^r(K,L) \otimes G \to \pi^r(K,L;G)$, and $\delta_{\#}$ defines

$$\beta\colon \pi^r(K,L;G) \to \operatorname{Tor}(\pi^{r+1}(K,L),G).$$

Hence the sequence (*) is exact under the hypothesis of 3.6. (*) is natural with respect to maps $f\colon (K,L) \to (K',L')$ because of 7.1 and the fact that the exact sequence of 3.5 is natural. A homomorphism $\phi\colon G \to H$ gives rise to a commutative diagram

$$\begin{array}{ccccccccc} 0 & \to & R & \xrightarrow{i} & F & \xrightarrow{j} & G & \to & 0 \\ & & \downarrow & & \downarrow & & \downarrow{\scriptstyle\phi} & & \\ 0 & \to & R' & \xrightarrow{i'} & F' & \xrightarrow{j'} & H & \to & 0, \end{array}$$

where R, R', F, and F' are finitely generated free abelian groups. Since R, R', F, F', and $G \varepsilon \mathfrak{D}$, by 3.5 we obtain a commutative diagram:

$$\begin{array}{ccccccc} \cdots \to \pi^r(K,L;G) & \xrightarrow{\delta_{\#}} & \pi^{r+1}(K,L;R) & \xrightarrow{i_{\#}} & \pi^{r+1}(K,L;F) \to \cdots \\ \downarrow{\scriptstyle\phi_{\#}} & & \downarrow & & \downarrow \\ \cdots \to \pi^r(K,L;H) & \xrightarrow{\delta_{\#}} & \pi^{r+1}(K,L;R') & \xrightarrow{i_{\#}} & \pi^{r+1}(K,L;F') \to \cdots. \end{array}$$

Hence the induced map $\operatorname{Tor}(\pi^{r+1}(K,L),G) \to \operatorname{Tor}(\pi^{r+1}(K,L),H)$ commutes with β. By 7.1, $\phi_{\#}$ commutes with α. This completes the proof of the first part of 3.6. Note the close similarity between this proof and the proof of Theorem 2.2A of [11].

An exact sequence $0 \to A \xrightarrow{i} B \xrightarrow{j} C \to 0$ *splits* if there exists a homomorphism $k\colon C \to B$ such that $jk =$ the identity on C. For abelian groups, this is equivalent to the statement that $B = A + C$, the direct sum of A and C.

LEMMA 7.2. *The exact sequence* (*) *splits for* $G = Z_p$, *p an odd prime.*

Proof. By 5.3, $\pi^r(K,L;Z_p)$ is a vector space over Z_p and hence, $\pi^r(K,L;Z_p)$, as a group, is a direct sum of copies of Z_p. This implies that (*) splits when $G = Z_p$.

LEMMA 7.3. *If* (K,L) *is such that* $\pi^r(K,L)$ *is finitely generated, then the exact sequence* (*) *splits for* $G = Z_{p^s}$, *p an odd prime.*

Proof. Corresponding to the coefficient homomorphism $\phi\colon Z_{p^s} \to Z_p$ sending a generator into a generator, we have the commutative diagram (by 7.1):

$$\begin{array}{ccc} \pi^r(K,L) \otimes Z_{p^s} & \xrightarrow{1\otimes\phi} & \pi^r(K,L) \otimes Z_p \\ \downarrow{\scriptstyle \alpha_1} & & \downarrow{\scriptstyle \alpha} \\ \pi^r(K,L;Z_{p^s}) & \xrightarrow{\phi_\#} & \pi^r(K,L;Z_p). \end{array}$$

α and α_1 are monomorphisms. We can write

$$\pi^r(K,L) = Z + Z + Z_{p_1^{m_1}} + \cdots + Z_{p_k^{m_k}}$$

by hypothesis. We obtain generators $\{x_j\}$ of $\pi^r(K,L) \otimes Z_{p^s}$ from generators of Z and $Z_{p_i^{m_i}}$ for $p_i = p$. It is obvious that $(1\otimes\phi)(x_j) \neq 0$ for all j, and hence $\alpha(1\otimes\phi)(x_j) \neq 0$ for all j. However, if (*) does not split for $G = Z_{p^s}$, then some x_j is such that $\alpha_1(x_j)$ is divisible by p in $\pi^r(K,L;Z_{p^s})$. Hence $\phi_\#\alpha_1(x_j)$ is divisible by p in $\pi^r(K,L;Z_p)$, but each element of $\pi^r(K,L;Z_p)$ has order p. Thus $0 = \phi_\#\alpha_1(x_j) = \alpha(1\otimes\phi)(x_j) \neq 0$. This is a contradiction.

LEMMA 7.4. (*H-lemma*). *Given a commutative diagram where each row and each column is exact. If the first and third columns and the middle row split, then every row and every column splits.*

$$\begin{array}{ccccccccc} & & 0 & & 0 & & 0 & & \\ & & \downarrow & & \downarrow & & \downarrow & & \\ 0 & \to & A_1 & \xrightarrow{f_1} & B_1 & \xrightarrow{g_1} & C_1 & \to & 0 \\ & & \downarrow{\scriptstyle c_1} & & \downarrow{\scriptstyle d_1} & & \downarrow{\scriptstyle e_1} & & \\ 0 & \to & A_2 & \xrightarrow{f_2} & B_2 & \xrightarrow{g_2} & C_2 & \to & 0 \\ & & \downarrow{\scriptstyle c_2} & & \downarrow{\scriptstyle d_2} & & \downarrow{\scriptstyle e_2} & & \\ 0 & \to & A_3 & \xrightarrow{f_3} & B_3 & \xrightarrow{g_3} & C_3 & \to & 0 \\ & & \downarrow & & \downarrow & & \downarrow & & \\ & & 0 & & 0 & & 0 & & \end{array}$$

Proof. By hypothesis, there are homomorphisms $c_2^{-1}\colon A_3 \to A_2$, $e_2^{-1}\colon C_3 \to C_2$, and $g_2^{-1}\colon C_2 \to B_2$ such that $c_2c_2^{-1} = 1$, $e_2e_2^{-1} = 1$, and $g_2g_2^{-1} = 1$ on A_3, C_3, and C_2 respectively. Define $g_3^{-1}\colon C_3 \to B_3$ by $g_3^{-1} = d_2g_2^{-1}e_2^{-1}$, then $g_3g_3^{-1} = g_3d_2g_2^{-1}e_2^{-1} = e_2g_2g_2^{-1}e_2^{-1} = e_2e_2^{-1} = 1$, and hence the bottom row splits. Let f_3^{-1} be the other component of the direct sum decomposition, i.e. $f_3^{-1}\colon B_3 \to A_3$ is such that $1 = f_3f_3^{-1} + g_3^{-1}g_3$ on B_3. Define $d_2^{-1} = g_2^{-1}e_2^{-1}g_3 + f_2c_2^{-1}f_3^{-1}\colon B_3 \to B_2$. Then

$$\begin{aligned} d_2d_2^{-1} &= d_2g_2^{-1}e_2^{-1}g_3 + d_2f_2c_2^{-1}f_3^{-1} = g_3^{-1}g_3 + f_3c_2c_2^{-1}f_3^{-1} \\ &= g_3^{-1}g_3 + f_3f_3^{-1} = 1, \end{aligned}$$

and hence the middle column splits. Let f_2^{-1} and c_1^{-1} be the other components of the direct sum decompositions of the second row and the first column, i.e. $f_2^{-1}f_2 = 1$ and $c_1^{-1}c_1 = 1$ on A_2 and A_1 respectively. Define $f_1^{-1}\colon B_1 \to A_1$ by $f_1^{-1} = c_1^{-1}f_2^{-1}d_1$. Then $f_1^{-1}f_1 = c_1^{-1}f_2^{-1}d_1f_1 = c_1^{-1}f_2^{-1}f_2c_1 = c_1^{-1}c_1 = 1$, and and hence every row and every column splits.

LEMMA 7.5. *If the exact sequence* (*) *splits for G' and G'', then it splits for $G' + G'' = G$, where G' and $G'' \varepsilon \mathfrak{D}$.*

Proof. $0 \to G' \to G \to G'' \to 0$ is a split exact sequence. The following is a commutative diagram:

$$\begin{array}{ccccc}
 & 0 & & 0 & & 0 \\
 & \downarrow & & \downarrow & & \downarrow \\
0 \to & \pi^r(K,L)\otimes G' & \to & \pi^r(K,L)\otimes G & \to & \pi^r(K,L)\otimes G'' \to 0 \\
 & \downarrow \alpha_1 & & \downarrow \alpha & & \downarrow \alpha_2 \\
0 \to & \pi^r(K,L;G') & \to & \pi^r(K,L;G) & \to & \pi^r(K,L;G'') \to 0 \\
 & \downarrow \beta_1 & & \downarrow \beta & & \downarrow \beta_2 \\
0 \to & \operatorname{Tor}(\pi^{r+1}(K,L),G') & \to & \operatorname{Tor}(\pi^{r+1}(K,L),G) & \to & \operatorname{Tor}(\pi^{r+1}(K,L),G'') \to 0 \\
 & \downarrow & & \downarrow & & \downarrow \\
 & 0 & & 0 & & 0
\end{array}$$

The rows are exact and split because $\otimes$, Tor, and $\pi^r(K,L;G)$ are additive functors when $G \varepsilon \mathfrak{D}$. The columns are exact by above, and the first and thirds columns split by hypothesis. Hence by 7.4, the middle column splits.

The last part of 3.6 now follows immediately from 7.3 and 7.5.

In order to prove 3.7, we first establish a direct limit theorem. Let $G \varepsilon \mathfrak{D}$, and let $\{G^\alpha; \pi_\alpha{}^\beta\}$ be the direct system of finitely generated subgroups of G [6; Chapt. VIII]. Then $G \approx \operatorname{dir}_\alpha \lim G^\alpha$. Note that each $G^\alpha \varepsilon \mathfrak{D}$. $\{\pi^r(K,L;G^\alpha); (\pi_\alpha{}^\beta)_\#\}$ is obviously a direct system of groups for $r > (N+1)/2$. We define $\xi\colon \operatorname{dir}_\alpha \lim \pi^r(K,L;G^\alpha) \to \pi^r(K,L;G)$ by

$$\xi(\{[a]\}) = (\pi_\alpha)_\#([a]),$$

where $[a] \varepsilon \pi^r(K,L;G^\alpha)$, $\{[a]\}$ is the element in $\operatorname{dir}_\alpha \lim \pi^r(K,L;G^\alpha)$ represented by $[a]$, and $\pi_\alpha\colon G^\alpha \to G$ is the canonical homomorphism. Since $\pi_r = \pi_\beta \pi_\alpha{}^\beta$ and all $G^\alpha \varepsilon \mathfrak{D}$, $(\pi_\alpha)_\# = (\pi_\beta)_\#(\pi_\alpha{}^\beta)_\#$ and ξ is well defined (note that $G \varepsilon \mathfrak{D}$ is necessary here).

THEOREM 7.6. *If (K,L) is a finite CW-pair, then ξ is an isomorphism for $r > (N+1)/2$. This isomorphism is natural with respect to maps $f\colon (K,L) \to (K',L')$ and homomorphisms $\phi\colon G \to H$.*

Proof. Let X be an $X(G,r)$-space, and let $[a] \varepsilon \pi^r(K,L;G)$. Since K is compact, $a(K) \subset X$ is also compact. Thus $a(K)$ is contained in a finite subcomplex X' of X [7; p. 96]. Assume X is the standard $X(G,r)$-space constructed in the proof of 3.1. Then a finite subcomplex of X must be an $X(G^\beta,r)$-space for some β. Hence a defines a map $a'\colon (K,L) \to (X',x_0)$ such that $(\pi_\beta)_\#([a']) = [a]$. Hence ξ is an epimorphism. A similar argument shows ξ is a monomorphism because $K \times I$ is also a compact CW-complex. The naturality statements are obvious.

Theorem 3.7 now follows immediately from the exact sequence (*), 7.6, and the fact that direct limits preserve exactness, $\otimes$, and Tor [5].

The following example shows that the universal coefficient theorem of 3.6 and 3.7 does not hold in general, even if $G \varepsilon \mathfrak{D}$. Let $G = Q =$ the additive group of the rationals and let K be an $X(Q,n)$-space. Then by 3.3, $\pi^n(K;Q) \approx \operatorname{Hom}(Q,Q) = Q$ and $\pi^n(K;Z) \approx \operatorname{Hom}(Q,Z) = 0$. The sequence (*), if exact, would give here that

$$0 \to 0 \otimes Q \to Q \to \operatorname{Tor}(\pi^{n+1}(K),Q) \to 0$$

is exact. However $\operatorname{Tor}(\pi^{n+1}(K),Q) = 0$, and hence $Q = 0$. This is a contradiction.

8. Proof of Theorem 3.8. The proof of Theorem 3.8, as the theorem itself, is a generalization of the proof of Theorem 3.1 of [11]. The proof is based on the generalized cohomotopy exact couple; i.e. Z is replaced everywhere by G.

Proof of 3.8. As is the proof of 3.1 of [11], it suffices to prove that

$$H^{r+1}(K,L;G_{(1)}) \varepsilon \mathscr{C}, \cdots, H^N(K,L;G_{(N-r)}) \varepsilon \mathscr{C}$$

for $r > \operatorname{Max}((N+1)/2, n - \alpha(\mathscr{C};G))$. Again $n - r < n - (n - \alpha(\mathscr{C};G)) = \alpha(\mathscr{C};G)$, hence $G_{(1)} \varepsilon \mathscr{C}, \cdots, G_{(n-r)} \varepsilon \mathscr{C}$ by definition of $\alpha(\mathscr{C};G)$. Since $G_{(s)}$ is finitely generated by 3.6, we may use the universal coefficient theorem for cohomology (2.3A of [11]); i.e.

$$0 \to H^{r+s}(K,L) \otimes G_{(s)} \xrightarrow{\alpha} H^{r+s}(K,L;G_{(s)}) \xrightarrow{\beta} \operatorname{Tor}(H^{r+s+1}(K,L),G_{(s)}) \to 0$$

is exact. $G_{(s)} \varepsilon \mathscr{C}$ for $s \leqq n - r$, therefore $H^{r+s}(K,L;G_{(s)}) \varepsilon \mathscr{C}$ for $s \leqq n - r$. The proof now differs from that of 3.1 of [11]. We have yet to show that

$H^{r+s}(K,L;G_{(s)}) \varepsilon \mathscr{C}$ for $s > n - r$ from the assumption that $H^{r+s}(K,L;G) \varepsilon \mathscr{C}$ for $s > n - r$. Since

$$(G+G')_{(s)} \approx G_{(s)} + G'_{(s)},\ H^r(K,L;G+G') \approx H^r(K,L;G) + H^r(K,L;G'),$$

and G is finitely generated, it suffices to show this for $G = Z$ or $G = Z_{p^t}$. The case $G = Z$ is obvious as in the proof of 3.1 of [11]. We now consider the case $G = Z_{p^t}$. By 3.6, $(Z_{p^t})_{(s)}$ can have only a p-primary component. Hence it suffices to show that if $H^r(K,L;Z_{p^t}) \varepsilon \mathscr{C}$ for $r > n$, then $H^r(K,L;Z_{p^s}) \varepsilon \mathscr{C}$ for $r > n$ and all $s \geqq 1$ (we have changed the notation to simplify the rest of the proof).

If $t = 1$, then it is obvious by induction on s using the exact sequence

$$H^r(K,L;Z_{p^{s-1}}) \to H^r(K,L;Z_{p^s}) \to H^r(K,L;Z_p)$$

corresponding to the exact coefficient sequence

$$0 \to Z_{p^{s-1}} \to Z_{p^s} \to Z_p \to 0.$$

If $t > 1$, then let r_0 be the largest integer r such that $H^r(K,L;Z_{p^q}) \not\varepsilon \mathscr{C}$ for some q with $1 \leqq q < t$, and assume $r_0 > n$. Then corresponding to the exact coefficient sequence

$$0 \to Z_{p^{t-q}} \to Z_{p^t} \to Z_{p^q} \to 0,$$

we have an exact sequence

$$H^{r_0}(K,L;Z_{p^t}) \to H^{r_0}(K,L;Z_{p^q}) \to H^{r_0+1}(K,L;Z_{p^{t-q}}).$$

However,

$$H^{r_0}(K,L;Z_{p^t}) \varepsilon \mathscr{C} \text{ and } H^{r_0+1}(K,L;Z_{p^{t-q}}) \varepsilon \mathscr{C},$$

hence $H^{r_0}(K,L;Z_{p^q}) \varepsilon \mathscr{C}$ which is a contradiction. Thus $r_0 \leqq n$ and $H^r(K,L;Z_p) \varepsilon \mathscr{C}$ for $r > n$ and as remarked above, this implies that $H^r(K,L;Z_{p^s}) \varepsilon \mathscr{C}$ for $r > n$ and all $s \geqq 1$. This completes the proof of 3.8.

9. Cohomotopy operations. We conclude our discussion of generalized cohomotopy groups by defining the concept of a universally defined cohomotopy operation analogous to universally defined homotopy operations [4]. We classify these operation and compute the classifying groups.

Let A and $B \varepsilon \mathfrak{D}$ throughout this section. A *cohomotopy operation* of type $(n,q;A,B)$ is a function $\theta\colon \pi^n(K,L;A) \to \pi^q(K,L;B)$, defined for every CW-pair (K,L) with $N \leqq \mathrm{Min}(2n-2, 2q-2)$, such that if $f\colon (K,L) \to (K',L')$, then $\theta f^\# = f^\# \theta\colon \pi^n(K',L';A) \to \pi^q(K,L;B)$.

Let $n+1 \leqq 2q-2$, and let $[b] \varepsilon \pi^q(X,x_0;B)$, where X is an $X(A,n)$-space. If $[a] \varepsilon \pi^n(K,L;A)$ and $N \leqq \mathrm{Min}(2n-2, 2q-2)$, then define

$\theta_b([a]) = [ba] \varepsilon \pi^q(K,L;B)$. Clearly θ_b is a cohomotopy operation of type $(n,q;A,B)$. Thus we have a function $\chi\colon \pi^q(X,x_0;B) \to$ the set of cohomotopy operations of type $(n,q;A,B)$, defined by $\chi([b]) = \theta_b$.

THEOREM 9.1. *If $n+1 \leqq 2q-2$, then χ is a 1-1 correspondence.*

Proof. Let $\iota \varepsilon \pi^n(X,x_0;A)$ denote the class of the identity map of (X,x_0) into (X,x_0). Let $a \varepsilon [a] \varepsilon \pi^n(K,L;A)$, $a\colon (K,L) \to (X,x_0)$. Then $a^\#(\iota) = [a]$. Hence if θ is a given cohomotopy operation of type $(n,q;A,B)$, then

$$\theta([a]) = \theta(a^\#(\iota)) = a^\#\theta(\iota) = \theta(\iota)[a] = \theta_b([a])$$

where $[b] = \theta(\iota) \varepsilon \pi^q(X,x_0;B)$. Thus χ is onto. If $\theta_b = \theta_{b'}$, then

$$[b] = [b]\iota = \theta_b(\iota) = \theta_{b'}(\iota) = [b']\iota = [b'].$$

Thus χ is 1-1.

COROLLARY 9.2. *If $n+1 \leqq 2q-2$, then each θ is a homomorphism.*

Proof. By 9.1, $\theta = \theta_b$ for some $[b] \varepsilon \pi^q(X,x_0;B)$. Let $[a]$ and $[a'] \varepsilon \pi^n(K,L;A)$, then

$$\theta_b([a]+[a']) = [b]([a]+[a']) = [ba]+[ba'] = \theta_b([a]) + \theta_b([a'])$$

by 2.3.

We now compute $\pi^q(X,x_0;B)$, where X is an $X(A,n)$-space.

THEOREM 9.3. (a) $\pi^n(X,x_0;B) \approx \mathrm{Hom}(A,B)$,

(b) $\pi^{n+1}(X,x_0;B) \approx \mathrm{Ext}(A,B)$,

(c) $\pi^q(X,x_0;B) = 0$ *for* $q > n+1$, *and*

(d) *if B is finitely generated and $q < n$, then*

$$0 \to \mathrm{Ext}(A, Z_{(n+1-q)} \otimes B + \mathrm{Tor}(Z_{(n-q)}, B)) \to \pi^q(X,x_0;B)$$
$$\to \mathrm{Hom}(A, Z_{(n-q)} \otimes B + \mathrm{Tor}(Z_{(n-q-1)}, B)) \to 0$$

is an exact sequence.

Proof. (a) follows from 3.3 and 3.4. (b) follows from 3.8 and the universal coefficient theorem for cohomology because X has dimension $n+1$. (c) is trivial because X has dimension $n+1$. If B is finitely generated, then $B_{(s)} \approx Z_{(s)} \otimes B + \mathrm{Tor}(Z_{(s-1)}, B)$ by Theorem 3.6.

$$
\begin{tikzcd}
 & & \vdots \arrow[d] & & & \\
 & \cdots \arrow[r] & 0 \arrow[r] \arrow[d, "j'"] & \cdots & \vdots \arrow[d] & \\
\cdots \arrow[r] & 0 \arrow[r, "\Delta'"] & \Gamma^{q,n} \arrow[r, "i'"] \arrow[d, "j'"] & H^{n+1}(X, x_0; B_{(n+1-q)}) \arrow[r, "\Delta'"] & 0 \arrow[r] \arrow[d] & \cdots \\
\cdots \arrow[r] & 0 \arrow[r, "\Delta'"] & \Gamma^{q,n-1} \arrow[r, "i'"] \arrow[d] & \mathrm{H}^{n}(X, x_0; B_{(n-q)}) \arrow[r, "\Delta'"] & 0 \arrow[r] \arrow[d] & \cdots \\
 & & \vdots \arrow[d] & & \vdots & \\
\cdots \arrow[r] & 0 \arrow[r, "\Delta'"] & \Gamma^{q,q} \arrow[r, "i'"] \arrow[d, "j'"] & 0 \arrow[r] & \cdots & \\
 & \cdots \arrow[r] & \pi^{q}(X, x_0; B) \arrow[r] \arrow[d] & 0 \arrow[r] & \cdots & \\
 & & \vdots & & &
\end{tikzcd}
$$

Figure 3.

Consider figure 3, a portion of the generalized cohomotopy exact couple with $G = B$, $(K, L) = (X, x_0)$, and $N = n + 1$. It is clear from figure 3 that the sequence

$$0 \to \Gamma^{q,n} \to \Gamma^{q,n-1} \to H^{n}(X, x_0; B_{(n-q)}) \to 0$$

is exact. However,

$$\Gamma^{q,n} \approx H^{n+1}(X, x_0; B_{(n+1-q)}) \approx \mathrm{Ext}(A, B_{(n+1-q)}), \Gamma^{q,n-1} \approx \Gamma^{q,q} \approx \pi^{q}(X, x_0; B),$$

and $H^{n}(X, x_0; B_{(n-q)}) \approx \mathrm{Hom}(A, B_{(n-q)})$. Combined with the above, this completes the proof.

10. Generalized homotopy groups. We conclude this paper with some brief remarks on generalized homotopy groups.

As one might expect, there is a theory of homotopy groups with coefficients in G which is dual in an intuitive sense to the theory developed above. The results of Spanier and Whitehead [17] make this duality precise. By these results, we are led to consider spaces having only one non-vanishing cohomology group; in particular, we consider homotopy classes of maps of such a space into arbitrary spaces. The duality of [17] gives theorems dual to our theorems on cohomotopy groups with coefficients in G. These theorems are only valid in the stable range. However, as in ordinary homotopy theory, there is a natural group structure defined outside of the stable range, and many of the theorems extend beyond the stable range. We leave the details of these results to the reader.

Appendix.

11. Proof of Lemma 3. 4. *Proof of* 3. 4.[5] For the properties of Ext needed in this proof, see [5]. There is a natural homomorphism

$$\chi\colon \operatorname{Hom}(A,B)\otimes Z_2 \to \operatorname{Hom}(A,B\otimes Z_2)$$

defined by $[\chi(f\otimes 1)](a)=f(a)\otimes 1$, where $f\,\varepsilon\operatorname{Hom}(A,B)$ and 1 is the non-zero element of Z_2. We first show that χ is an isomorphism if A is free. Let $A=\sum_i Z_i, i\,\varepsilon\, I$. Then

$$\operatorname{Hom}(A,B)=\operatorname{Hom}(\sum_i Z_i,B)\approx\prod_i\operatorname{Hom}(Z_i,B)=\prod_i B_i$$

and

$$\operatorname{Hom}(A,B\otimes Z_2)=\operatorname{Hom}(\sum_i Z_i,B\otimes Z_2)\approx\prod_i\operatorname{Hom}(Z_i,B\otimes Z_2)=\prod_i(B\otimes Z_2)_i.$$

Moreover, the natural homomorphism $(\prod_i B_i)\otimes Z_2\to\prod_i(B\otimes Z_2)_i$ is an isomorphism (see exercise E-6 in [6; Chapt. V]). Hence χ is an isomorphism if A is free.

Let $0\to R\to F\to G\to 0$ be exact, where F and R are free abelian groups. By definition of Ext, $\operatorname{Hom}(F,H)\to\operatorname{Hom}(R,H)\to\operatorname{Ext}(G,H)\to 0$ is exact. Thus the following is a commutative diagram with the rows exact:

$$\begin{array}{ccccccccc}
\operatorname{Hom}(F,H)\otimes Z_2 & \to & \operatorname{Hom}(R,H)\otimes Z_2 & \to & \operatorname{Ext}(G,H)\otimes Z_2 & \to & 0 & \to & 0\\
\downarrow & & \downarrow & & \downarrow & & \downarrow & & \downarrow\\
\operatorname{Hom}(F,H\otimes Z_2) & \to & \operatorname{Hom}(R,H\otimes Z_2) & \to & \operatorname{Ext}(G,H\otimes Z_2) & \to & 0 & \to & 0.
\end{array}$$

The first two and the last two vertical homomorphisms are isomorphisms, and hence by the five-lemma [6; p. 16], $\operatorname{Ext}(G,H)\otimes Z_2\approx\operatorname{Ext}(G,H\otimes Z_2)$. Now by hypothesis, $0\to G\xrightarrow{\xi} G$ is exact, where $\xi(g)=2g$. Thus

$$\operatorname{Ext}(G,H)\xrightarrow{\xi^*}\operatorname{Ext}(G,H)\to 0$$

is exact, where ξ^* is multiplication by 2. Hence every element of $\operatorname{Ext}(G,H)$ is divisible by 2 and $\operatorname{Ext}(G,H\otimes Z_2)\approx\operatorname{Ext}(G,H)\otimes Z_2=0$. This completes the proof.

PRINCETON UNIVERSITY.

---

[5] This proof was worked out with the help of D. A. Buchsbaum.

REFERENCES.

[1] M. G. Barratt, "Track groups I," *Proceedings of the London Mathematical Society*, vol. 5 (1955), pp. 71-106.

[2] ———, "Homotopy ringoids and homotopy groups," *Quarterly Journal of Mathematics*, vol. 5 (1954), pp. 271-290.

[3] A. L. Blakers and W. S. Massey, "The homotopy groups of a triad, I," *Annals of Mathematics*, vol. 53 (1951), pp. 161-205.

[4] ———, "Products in homotopy theory," *Annals of Mathematics*, vol. 58 (1953), pp. 295-324.

[5] H. Cartan and S. Eilenberg, *Homological algebra*, Princeton University Press, 1956.

[6] S. Eilenberg and N. E. Steenrod, *Foundations of algebraic topology*, Princeton University Press, 1953.

[7] P. J. Hilton, *An introduction to homotopy theory*, Cambridge University Press, 1953.

[8] W. S. Massey, "Exact couples in algebraic topology," I and II, III, IV, and V, *Annals of Mathematics*, vol. 56 (1952), pp. 363-396, vol. 57 (1953), pp. 248-286.

[9] J. C. Moore, "Some applications of homology theory to homotopy problems," *Annals of Mathematics*, vol. 58 (1953), pp. 325-350.

[10] ———, "On homotopy groups of spaces with a single non-vanishing homology group," *Annals of Mathematics*, vol. 59 (1954), pp. 549-557.

[11] F. P. Peterson, "Some results on cohomotopy groups," *American Journal of Mathematics*, vol. 78 (1956), pp. 243-258.

[12] J.-P. Serre, "Homologie singuliere des espaces fibrés," *Annals of Mathematics*, vol. 54 (1951), pp. 425-505.

[13] ———, "Groupes d'homotopie et classes de groupes abéliens," *Annals of Mathematics*, vol. 58 (1953), pp. 258-294.

[14] E. H. Spanier, "Borsuk's cohomotopy groups," *Annals of Mathematics*, vol. 50 (1949), pp. 203-245.

[15] ——— and J. H. C. Whitehead, "A first approximation to homotopy theory," *Proceedings of the National Academy of Sciences*, vol. 39 (1953), pp. 655-660.

[16] ——— and J. H. C. Whitehead, "The theory of carriers and S-theory," *Algebraic Geometry and Topology* (A Symposium in Honor of S. Lefschetz), Princeton University Press, 1956.

[17] ——— and J. H. C. Whitehead, "Duality in homotopy theory," *Mathematika*, vol. 2 (1955), pp. 56-80.

[18] J. H. C. Whitehead, "Combinatorial homotopy," I, *Bulletin of the American Mathematical Society*, vol. 55 (1949), pp. 213-245.

A NOTE ON THE INTERPOLATION OF SUBLINEAR OPERATIONS.*

By A. P. CALDERÓN and A. ZYGMUND.[1]

The purpose of this note is to give an extension of M. Riesz' interpolation theorem for linear operations to certain non-linear ones.

Let R be a measure space. This means that we have a set function $\mu(E)$, non-negative and countably additive, defined for some ('measurable') subsets E of R. For any measurable (with respect to μ) function f defined on R we write

$$\left(\int_R |f|^r d\mu\right)^{1/r} = \| f \|_{r,\mu} \qquad (0 < r < \infty),$$

and denote by $\| f \|_{\infty,\mu}$ the essential (with respect to μ) upper bound of $|f|$. The set of functions f such that $\| f \|_{r,\mu}$ is finite $(0 < r \leqq \infty)$ is denoted by $L^{r,\mu}$. If no confusion arises, we write $\| f \|_r$, L^r for $\| f \|_{r,\mu}$, $L^{r,\mu}$.

Let R_1 and R_2 be two measure spaces with measures μ and ν respectively. Let $h = Tf$ be a transformation of functions $f = f(u)$ defined (almost everywhere) on R_1 into functions $h = h(v)$ defined on R_2. The most important special case is when T is a *linear operation*. This means that if Tf_1 and Tf_2 are defined, and if α_1, α_2 are complex numbers, then $T(\alpha_1 f_1 + \alpha_2 f_2)$ is defined and

$$T(\alpha_1 f_1 + \alpha_2 f_2) = \alpha_1 Tf_1 + \alpha_2 Tf_2.$$

Let $r > 0$, $s > 0$. A linear operation $h = Tf$ will be said to be *of type* (r, s) if it is defined for each $f \varepsilon L^{r,\mu}$ and if

$$\| Tf \|_{s,\nu} \leqq M \| f \|_{r,\mu}, \tag{1}$$

where M is independent of f. The least value of M is called the (r, s) *norm* of the operation.

Denote by (α, β) points of the square

$$0 \leqq \alpha \leqq 1, \qquad 0 \leqq \beta \leqq 1. \tag{Q}$$

---

* Received September 9, 1955.

[1] The research resulting in this paper was supported in part by the office of Scientific Research of the Air Force under contract AF 18(600)-1111.

The Riesz interpolation theorem (in the form generalized by Thorin (see [1]-[6] of the References at the end of the note) asserts that if a linear operation $h = Tf$ is simultaneously of types $(1/\alpha_1, 1/\beta_1)$ and $(1/\alpha_2, 1/\beta_2)$, with norms M_1 and M_2 respectively, and if

$$(2) \qquad \alpha = (1-t)\alpha_1 + t\alpha_2, \quad \beta = (1-t)\beta_1 + t\beta_2, \qquad (0 < t < 1)$$

then T is also of type $(1/\alpha, 1/\beta)$, with norm

$$(3) \qquad M \leqq M_1^{1-t} M_2^{t}.$$

The significance of this theorem is by now widely recognized, and its applications are many. Riesz himself deduced the result, through appropriate passages to limits, from a theorem about bilinear forms, and in this argument the linearity of T plays an important role. The same can be said of other proofs. There are however a number of interesting operations which are not linear and to which therefore the theorem cannot be applied. For the sake of illustration we mention one of them, first considered by Littlewood and Paley (see [7]), which has important application in Fourier series.

Given any $f \varepsilon L(0, 2\pi)$, we consider the function $F(z)$ regular for $|z| < 1$, whose real part is the Poisson integral of f, and imaginary part is zero at the origin. The Littlewood-Paley function is

$$g(\theta) = \{ \int_0^1 (1-\rho) |F'(\rho e^{i\theta})|^2 d\rho \}^{\frac{1}{2}}.$$

The operation $g = Tf$ is clearly not linear. It satisfies, however, the following relations

$$(4) \qquad |T(f_1 + f_2)| \leqq |Tf_1| + |Tf_2|,$$

$$(5) \qquad |T(kf)| = |k| \, |Tf|,$$

for any constant k.

There are other interesting non-linear operations which have the same properties and it may be of interest to study the problem of interpolation of such operations. This is the object of this note.

We begin with general definitions.

We call an operation $h = Tf$ *sublinear*, if the following conditions are satisfied:

(i) Tf is defined (uniquely) if $f = f_1 + f_2$, and Tf_1 and Tf_2 are defined;

(ii) For any constant k, $T(kf)$ is defined if Tf is defined;

(iii) Conditions (4) and (5) hold.

In view of (5) we may, as in the linear case, consider inequalities (1) and introduce the notions of the *type* and *norm* of a sublinear operation. In what follows, the functions f will be defined (almost everywhere) on a space R_1 with measure μ, and the $h = Tf$ on a space R_2 with measure ν.

THEOREM. *Let (α_1, β_1) and (α_2, β_2) be any two points of the square Q. Suppose that a sublinear operation $h = Tf$ is simultaneously of types $(1/\alpha_1, 1/\beta_1)$ and $(1/\alpha_2, 1/\beta_2)$ with norms M_1 and M_2 respectively. Let (α, β) be given by (2). Then T is also of type $(1/\alpha, 1/\beta)$, with norm M satisfying (3).*

Proof. We easily deduce from conditions (i), (ii), (iii) that, if Tf_1, $Tf_2, \cdots, Tf_n$ are defined so is $T\{n^{-1}(f_1 + \cdots + f_n)\}$ and

$$(6) \qquad | T\{(f_1 + f_2 + \cdots + f_n)/n\} | \leqq n^{-1}(| Tf_1 | + \cdots + | Tf_n |).$$

We may suppose that $\alpha_1 \leqq \alpha_2$. Thus

$$(7) \qquad \alpha_1 \leqq \alpha \leqq \alpha_2.$$

Consider any f in $L^{1/\alpha, \mu}$ and write $f = f_1 + f_2$, where f_1 equals f at the points at which $|f| \leqq 1$, and equals 0 elsewhere. By (7),

$$| f_1 |^{1/\alpha_1} \leqq | f_1 |^{1/\alpha} \leqq | f |^{1/\alpha},$$

so that $f_1 \varepsilon L^{1/\alpha_1}$ and Tf_1 is defined, by hypothesis. Similarly $f_2 \varepsilon L^{1/\alpha_2}$ and Tf_2 is defined. It follows from (i) that $Tf = T(f_1 + f_2)$ is defined. Our task is to show that the $(1/\alpha, 1/\beta)$ norm M of T is finite and satisfies (3).

We assume for the time being that $\alpha > 0$, $\beta < 1$. Take any $f \varepsilon L^{1/\alpha}$. Without loss of generality we may suppose that

$$\| f \|_{1/\alpha} = 1.$$

Clearly

$$(8) \qquad \| Tf \|_{1/\beta} = \sup_g \int_{R_2} | Tf | \cdot g d\nu,$$

where g is non-negative and satisfies $\| g \|_{1/(1-\beta)} = 1$. We may confine our attention to functions g which are simple (a function is called *simple* if it takes only a finite number of values and is distinct from 0 on a set of finite measure; simple functions are dense in every L^s, $0 < s < \infty$; and in L^∞, if the space has finite measure). We make one more assumption, of which we shall free ourselves later, namely that f is also simple. We fix f and g and consider the integral

$$(9) \qquad I = \int_{R_2} | Tf | \cdot g d\nu.$$

Let $c_1, c_2, \cdots, c_m$ be the distinct from 0 (and different from each other) values of f. Let E_k be the set in which $f = c_k$, and let $\chi_k = \chi_k(u)$ be the characteristic function of E_k. Similarly let $c'_1, c'_2, \cdots, c'_n$ be the different from 0 values of g, E'_l the set where $g = c'_l$, and $\chi_l = \chi_l(v)$ the characteristic function of E'_l. Hence $f = \sum | c_k | \epsilon_k \chi_k$, $g = \sum c'_l \chi'_l$, where $| \epsilon_k | = 1$ and $c'_l > 0$.

Let $\alpha(z)$ and $\beta(z)$ be the right sides of (2), with z for t. Consider the non-negative function

$$\Phi(z) = \int_{R_2} | T(| f |^{\alpha(z)/\alpha} \operatorname{sign} f) | \, g^{(1-\beta(z))/(1-\beta)} \, d\nu \qquad (z = x + iy), \tag{10}$$

which reduces to I for $z = t$. We show that $\Phi(z)$ is continuous and $\log \Phi(z)$ is subharmonic, in the whole plane.

Since, for each z, $| f |^{\alpha(z)/\alpha} \operatorname{sign} f$ is simple, and so is in L^{1/α_1}, the function $T(| f |^{\alpha(t)/\alpha} \operatorname{sign} f)$ is in L^{1/β_1}, and in particular is integrable over the set where $g > 0$. Hence $\Phi(z)$ exists for each z.

We have

$$\begin{aligned} \Phi(z) &= \int_{R_2} | T\{\sum | c_k |^{\alpha(z)/\alpha} \epsilon_k \chi_k\} | \, \{\sum c'_l{}^{(1-\beta(x))(1-\beta)} \chi_l\} \, d\nu \\ &= \sum \int_{E'_l} c'_l{}^{(1-\beta(x))/(1-\beta)} \, | T\{ \sum | c_k |^{\alpha(z)/\alpha} \epsilon_k \chi_k\} | \, d\nu \\ &= \sum \int_{E'_l} T\{c'_l{}^{(1-\beta(x))/(1-\beta)} \sum | c_k |^{\alpha(z)/\alpha} \epsilon_k \chi_k\} | \, d\nu, \end{aligned} \tag{11}$$

and it is enough to show that each integral of the last sum is continuous and its logarithm is subharmonic. In proving this we shall make repeated use of the inequality $| \, | Tf_1 | - | Tf_2 | \, | \leqq | T(f_1 - f_2) |$, which is a consequence of (4).

We therefore fix l and write

$$\psi_z = \sum_k c'_l{}^{(1-\beta(z))/(1-\beta)} | c_k |^{\alpha(z)/\alpha} \epsilon_k \chi_k, \qquad \Psi(z) = \int_{E'_l} | T\psi_z | \, d\nu.$$

Clearly

$$\begin{aligned} | \Psi(z + \Delta z) - \Psi(z) | &\leqq \int_{E'_l} | T(\psi_{z+\Delta z} - \psi_z) | \, d\nu \\ &\leqq \| T(\psi_{z+\Delta z} - \psi_z) \|_{1/\beta_1} \{\nu(E'_l)\}^{1-\beta_1} \\ &\leqq M_1 \| \psi_{z+\Delta z} - \psi_z \|_{1/\alpha_1} \{\nu(E'_l)\}^{1-\beta_1}, \end{aligned}$$

and since $\psi_{z+\Delta z} - \psi_z$ is zero outside $\cup E_k$ and tends to 0, uniformly in u, as $\Delta z \to 0$, the norm $\| \psi_{z+\Delta z} - \psi_z \|_{1/\alpha_1}$ tends to 0, and Ψ is continuous at z. Hence Φ is continuous.

It is very well known that $\log\Psi(z)$ is subharmonic if and only if $\Psi(z)e^{h(z)}$ is subharmonic for every harmonic $h(z)$. We fix a harmonic function $h(z)$, and denote by $H(z)$ the analytic function whose real part is $h(z)$. Since the problem is local, we may consider h and H in a given circle. Write

$$\psi^*_z = \psi_z e^{H(z)}, \qquad \Psi^*(z) = \Psi(z)e^{h(z)} = \int_{E'_l} | T\psi^*_z | \, d\nu.$$

We fix z, take a $\rho > 0$, and denote by $z_1, z_2, \cdots, z_p$ a system of points equally spaced over the circumference of the circle with center z and radius ρ. We have

$$\psi^*_z(u) = \lim_{p\to\infty} 1/p \sum_{j=1}^{p} \psi^*_{z_j}(u),$$

uniformly in u. Since

$$\int_{E'_l} | T(\psi^*_z - 1/p \sum_1^p \psi_{z_j}) | \, d\nu \leqq \| T(\psi^*_z - 1/p \sum_1^p \psi^*_{z_j} \|_{1/\beta_1} \{\nu(E'_l)\}^{1-\beta_1}$$
$$\leqq M_1 \| \psi^*_z - 1/p \sum_1^p \psi^*_{z_j} \|_{1/\alpha_1} \{\nu(E'_l\}^{1-\beta_1}$$

the left side tends to 0 as $p \to \infty$. In particular, as $p \to \infty$ we have

$$\delta_p = \int_{E'_l} | \, | T\psi^*_z | - | T(1/p \sum_1^p \psi^*_{z_j}) | \, | \, d\nu \to 0,$$

$$\int_{E'_l} | T\psi^*_z | \, d\nu \leqq \delta_p + 1/p \sum_1^p \int_{E'_l} | T\psi^*_{z_j} | \, d\nu,$$

$$\Psi^*(z) \leqq \lim_{p\to\infty} 1/p \sum_1^p \Psi^*(z_j) = 1/(2\pi) \int_0^{2\pi} \Psi^*(z + \rho e^{it}) \, dt.$$

Hence $\Psi^*(z)$ is subharmonic.

We have therefore proved that $\Phi(z)$ is continuous in the whole plane and its logarithm is subharmonic. Moreover $\Phi(z)$ is bounded in every vertical strip of the plane, since from (11), (4) and (5) we deduce that

$$\Phi(z) \leqq \sum_{k,l} | c_k |^{\alpha(z)/\alpha} c'_l{}^{(1-\beta(x))/(1-\beta)} \int_{E'_l} | T(\chi_k) | \, d\nu.$$

Next we show that $\Phi \leqq M_1$ on the line $x = 0$, and $\Phi \leqq M_2$ on $x = 1$. It is enough to prove the first inequality. If $x = 0$, then

$$\Phi(z) \leqq \| T(|f|^{\alpha(z)/\alpha} \operatorname{sign} f) \|_{1/\beta_1} \| g^{(1-\beta_1)/(1-\beta)} \|_{1/(1-\beta_1)}$$
$$\leqq M_1 \| \, |f|^{\alpha(z)/\alpha} \operatorname{sign} f \|_{1/\alpha_1} \leqq M_1 \| \, |f|^{\alpha_1/\alpha} \|_{1/\alpha_1} = M_1.$$

Since $\log\Phi(z)$ is bounded above and subharmonic in the strip $0 \leqq x \leqq 1$,

and does not exceed $\log M_1$ and $\log M_2$ on the lines $x=0$ and $x=1$ respectively, an application of the Three-Line Theorem for subharmonic functions shows that

$$\log \Phi(z) \leqq (1-t) \log M_1 + t \log M_2$$

on the line $x=t$ and, in particular, $I=\Phi(t) \leqq M_1^{1-t} M_2^{t}$.

Summarizing results, we have proved that

$$\| Tf \|_{1/\beta} \leqq M_1^{1-t} M_2^{t} \| f \|_{1/\alpha} \tag{12}$$

for each simple f. We show that this holds for every $f \varepsilon L^{1/\alpha}$.

We fix such an f, and for each $m=1,2,\cdots$ consider the decomposition $f=f'_m+f''_m$, in which $f'_m=f$ wherever $|f| \leqq m$, and $f'_m=0$ elsewhere; hence $|f''_m|$ is either 0 or else greater than m. Let f_m be a simple function equal to 0 wherever $f'_m=0$ and such that $|f_m-f'_m|<1/m$ everywhere. Then

$$| \, |Tf|-|Tf_m| \, | \leqq |T(f-f_m)| \leqq |T(f'_m-f_m)|+|Tf''_m|. \tag{13}$$

If we show that each term on the right tends to 0 almost everywhere as m tends to $+\infty$ through a sequence of values, then the inequality (12) with f_m for f, will lead, by Fatou's lemma, to the inequality for f.

Now f'_m-f_m is in $L^{1/\alpha}$, and so also in L^{1/α_1}, since $|f'_m-f_m|<1/m$. It follows that

$$\| T(f'_m-f_m) \|_{1/\beta_1} \leqq M_1 \| f'_m-f_m \|_{1/\alpha_1} \leqq M_1 \| f'_m-f_m \|_{1/\alpha}^{\alpha_1/\alpha} \to 0. \tag{14}$$

as $m \to \infty$. Similarly

$$\| Tf''_m \|_{1/\beta_2} \leqq M_2 \| f''_m \|_{1/\alpha_2} \leqq M_2 \| f''_m \|_{1/\alpha}^{\alpha_2/\alpha} \to 0. \tag{15}$$

The inequalities (14) and (15) imply that there is a sequence of m tending to $+\infty$ and such that $|T(f'_m-f_m)|$ and $|Tf''_m|$ tend to 0 almost everywhere. This completes the proof of the theorem.

It remains however to consider the two extreme cases $\alpha=0$ and $\beta=1$, which we previously put aside. These two cases cannot occur simultaneously. If $\beta=1$, we replace the right side of (8) by $\int_{R_2} |Tf| \, d\nu$ and the function $\Phi(z)$ of (10) by

$$\int_{R_2} | T(|f|^{\alpha(z)/\alpha} \operatorname{sign} f) | \, d\nu.$$

After that the proof proceeds as before. If $\alpha=0$, then also $\alpha_1=\alpha_2=0$;

but it is immediately seen that whenever $\alpha_1 = \alpha_2$ the theorem is a corollary of Hölder's inequality.

MASSACHUSETTS INSTITUTE OF TECHNOLOGY AND
THE UNIVERSITY OF CHICAGO.

REFERENCES.

[1] M. Riesz, "Sur les maxima des formes billinéaires et sur les fonctionnelles linéaires," *Acta Mathematica*, vol. 49 (1926), pp. 465-497.

[2] G. O. Thorin, "An extension of a convexity theorem due to M. Riesz," *Kungl. Fysiografiska Saellskapets i Lund Forhaendliger*, 8 (1939), Nr. 14.

[3] ———, *Convexity theorems*, Uppsala, 1948, pp. 1-57.

[4] R. Salem, "Convexity theorems," *Bulletin of the American Mathematical Society*, vol. 55 (1949), pp. 851-860.

[5] J. D. Tamarkin and A. Zygmund, "Proof of a theorem of Thorin," *Bulletin of the American Mathematical Society*, vol. 50 (1944), pp. 279-282.

[6] A. P. Calderón and A. Zygmund, "On the theorem of Hausdorff-Young," *Annals of Mathematical Studies*, vol. 25 (1950), pp. 166-188.

[7] J. E. Littlewood and R. E. A. C. Paley, "Theorems on Fourier series and power series, I," *Journal of the London Mathematical Society*, vol. 6 (1931), pp. 230-233.

ON SINGULAR INTEGRALS.*

By A. P. Calderón and A. Zygmund.[1]

1. Introduction. In earlier work [1] we considered certain singular integrals arising in various problems of Analysis and studied some of their properties. Here we present a new approach to such integrals. Unlike the method used in [1] it is based on the theory of Hilbert transforms of functions of one variable, but otherwise it is simpler and yields most results obtained previously, under far less restrictive assumptions. Unfortunately some important cases ($f \varepsilon L$ for instance) seem to be beyond its scope. We have been unable to decide whether the corresponding theorems as presented in [1] can be likewise strengthened.

Our present results can be summarized in the theorems presented below.

Let $x, y, z, \cdots$ denote vectors in n-dimensional Euclidean space E_n, $|x|$ the length of x and $x' = x|x|^{-1}$. Consider the integral

$$1.1 \qquad \tilde{f}_\epsilon(x) = \int_{|x-y|>\epsilon} K(x,y) f(y)\, dy,$$

where dy denotes the element of volume in E_n.

Theorem 1. *If $K(x,y) = N(x-y)$, where $N(x)$ is a homogeneous function of degree $-n$, i.e. such that $N(\lambda x) = \lambda^{-n} N(x)$ for every x and $\lambda > 0$, and if $N(x)$ has in addition the following properties*

i) *$N(x)$ is integrable over the sphere $|x| = 1$ and its integral is zero,*

ii) *$N(x) + N(-x)$ belongs to $L \log^+ L$ on $|x| = 1$,*

then, if $f(x) \varepsilon L^p$, $1 < p < \infty$, $\tilde{f}_\epsilon(x)$ as defined in 1.1 converges to a limit $\tilde{f}(x)$ in the mean of order p, and pointwise almost everywhere as $\epsilon \to 0$. Furthermore $\bar{f}(x) = \sup_\epsilon |\tilde{f}_\epsilon(x)|$ belongs to L^p and $\|\bar{f}\|_p \leqq A \|f\|_p$, where A is a constant depending on p and K, and $\|f\|_p$ is the L^p norm of f.

The condition that $N(x) + N(-x)$ be in $L \log^+ L$ on $|x| = 1$ cannot be relaxed. For given a function $\phi(t)$ such that $\phi(t)/t \log t \to 0$ as $t \to \infty$

* Received September 9, 1955.

[1] The research resulting in this paper was supported in part by the office of Scientific Research of the Air Force under contract AF 18(600)-1111.

there exists a function satisfying i) such that $\phi[|N(x)+N(-x)|]$ is integrable on $|x|=1$ but whose Fourier transform is unbounded, so that even if the pointwise limit of $\tilde{f}_\epsilon(x)$ exists (as is the case of $f(x)$ continuously differentiable and vanishing outside a bounded set), no relationship of the form $\|\tilde{f}\|_2 \leqq A\|f\|_2$ holds.

The Fourier transform $M(x)$ of $N(x)$ is a homogeneous function of degree zero and can easily be shown to be given by the formula

$$M(x) = \int N(y')\,[i\tfrac{1}{2}\pi \operatorname{sg}\cos\theta + \log|\cos\theta|]\,d\sigma,$$

where θ is the angle between the unit vectors x' and y', and $d\sigma$ is the element of "area" of the sphere $|x|=1$ over which the integral is extended. It is the presence of the term $\log|\cos\theta|$ which makes the class $L\log^+L$ the best possible. Since we merely want to indicate this fact we omit further details.

THEOREM 2. *If*

$$K(x,y) = N(x, x-y)$$

where $N(x,y)$ is homogeneous of degree $-n$ in y and

i) *for every x, $N(x,y)$ is integrable over the sphere $|y|=1$ and its integral is zero,*

ii) *for a $q>1$ and every x, $|N(x,y)|^q$ is integrable over the sphere $|y|=1$ and its integral is bounded,*

then the same conclusions as in Theorem 1 hold about $\tilde{f}_\epsilon(x)$, provided that $f \in L^p$ with $q/(q-1) \leqq p < \infty$.

The condition that $p \geqq q/(q-1)$ is essential. We shall show by means of an example that if $p < q/(q-1)$, then $\tilde{f}_\epsilon(x)$ need no longer be in L^p.

A third type of integrals suggested by the theory of spherical summability of Fourier integrals is the object of the next two theorems.

THEOREM 3. *If*

$$K(x,y) = N(x, x-y)\psi(|x-y|)$$

where $\psi(t)$ is a Fourier-Stieltjes transform, $N(x,y)$ is homogeneous of degree $-n$ in y and

i) *$|N(x,y)| \leqq F(y)$ where $F(y)$ is a homogeneous function of degree $-n$ integrable over $|y|=1$,*

ii) *$\psi(t)$ is an even function and $N(x,y)$ is odd in y, i.e. $N(x,y) = -N(x,-y)$, or $\psi(t)$ is odd and $N(x,y)$ is even in y,*

then the same conclusions as in Theorem 1 hold about $\tilde{f}_\epsilon(x)$.

THEOREM 4. *If $K(x,y)$ is the same as in the previous theorem with condition* i) *replaced by*

i′) *for some $q > 1$ and every x, $|N(x,y)|^q$ is integrable over the sphere $|y| = 1$ and its integral is bounded,*

then the same conclusions about $\tilde{f}_\epsilon(x)$ hold provided that $f \in L^p$, $q/(q-1) \leqq p < \infty$.

In the cases of Theorems 2, 3 and 4 we may also consider the transposed integral 1.1, that is $\int_{|x-y|>\epsilon} K(y,x)f(y)dy$. The convergence in the mean of this integral in an immediate consequence of those theorems. The pointwise convergence does not follow readily though, and at individual points the integral may actually diverge even if f is continuously differentiable and vanishes outside a bounded set.

A straightforward application of Theorem 3 will yield the following statement about spherical summability of Fourier integrals.

THEOREM 5. *If the number n of variables of f is odd and $f \in L^p$, $1 < p \leqq 2$, then the spherical means of order $\frac{1}{2}(n-1)$ of the Fourier integral representation of f converge to f in the mean of order p.*

Whether this theorem remains valid for n even is an open question.

Finally, we might also mention two generalizations of the maximal theorem of Hardy and Littlewood which are obtained using the same ideas. These extensions are needed in the proofs of Theorems 1 and 2.

THEOREM 6. *Let $K_\epsilon(x,y) = \epsilon^{-n}N(x-y)\psi(\epsilon^{-1}|x-y|)$ where $N(x)$ is a non-negative homogeneous function of degree zero, integrable over $|x| = 1$, and $\psi(t)$ is a non increasing function of the real variable t such that $\psi(|x|)$ is integrable in E_n. Then if $f \in L^p$, $1 < p \leqq \infty$, and*

$$f^*(x) = \sup_\epsilon \left| \int K_\epsilon(x,y)f(y)dy \right|,$$

f^* belongs to L^p and

$$\| f^* \|_p \leqq A \| f \|_p,$$

where A is a constant depending on N, p and ψ.

The case when $N(x) \equiv 1$ and $x(t)$ is the characteristic function of the interval $(0,1)$ is well known.

THEOREM 7. *If* $K_\epsilon(x,y) = \epsilon^{-n} N(x, x-y) \psi(\epsilon^{-1} | x-y |)$ *where* $N(x,y)$ *is homogeneous of degree zero in* y, $| N(x,y) |^q$, $q > 1$, *is integrable over the sphere* $| y | = 1$ *and its integral is bounded, and* $\psi(t)$ *is the same as in the previous theorem, then* f^* *as defined in Theorem* 6 *is in the same* L^p *class as* f *and* $\| f^* \|_p \leqq A \| f \|_p$, *provided that* $q/(q-1) \leqq p < \infty$.

2. We start by showing that the integral 1.1 is meaningful. In the cases of Theorems 1 and 3 this is not quite evident.

In either case we have $| K(x,y) | \leqq F(| x-y |)$, where $F(y)$ is a homogeneous function of degree $-n$, integrable over the sphere $| y | = 1$, and thus it will be sufficient to show that

$$\int_{|x-y|>\epsilon} F(x-y) \, | f(y) | \, dy$$

is absolutely convergent for almost every x and any $\epsilon > 0$.

Let y' be a unit vector, t a real number and S a full sphere in E_n of diameter d. Then the integral

2.1 $$\int_\Sigma F(y') dy' \int_S dx \int_\epsilon^\infty | t^{-1} f(x - ty') | \, dt,$$

where dy' is the element of area of the unit sphere Σ, is finite. In fact, the inner integral is less than or equal to

$$A \Big[\int_{-\infty}^{+\infty} | f(x - ty') |^p \, dt \Big]^{1/p},$$[2]

where A depends on ϵ and p but not on f. Substituting this expression for the inner integral in 2.1 and applying Hölder's inequality to the integral over S we find that the latter is dominated by

$$A \, | S |^{(p-1)/p} \Big[\int_S dx \int_{-\infty}^{+\infty} | f(x - ty') |^p \, dt \Big]^{1/p}.$$

Now the integral with respect to x can be computed first along lines parallel to y' and then over the space of such lines, rendering evident that its value does not exceed $d \| f \|_p^p$. Thus the integral over S in 2.1 is a bounded function of y' and 2.1 is therefore finite. Hence

$$\int_\Sigma F(y') dy' \int_\epsilon^\infty | t^{-1} f(x - ty') | \, dt$$

[2] Throughout the rest of the paper the letter A will stand for a constant, not necessarily the same in each occurrence.

is finite for almost all x. But the last is nothing but the expression of $\int_{|x-y|>\epsilon} F(x-y)\,|f(y)|\,dy$ in polar coordinates with origin at x. In other words, for any $\epsilon > 0$, 1.1 is absolutely convergent for almost every x.

3. In this section we shall prove Theorems 3 and 4.

Let $g(t)$ be a function of the real variable t belonging to L^p in $-\infty < t < \infty$, $1 < p < \infty$, and let $\epsilon(s)$ be an arbitrary positive measurable function in $-\infty < s < \infty$. Then the integral

$$\int_{|s-t|>\epsilon(s)} g(t)/(s-t)\,dt$$

represents a function whose L^p norm does not exceed the L^p norm of g multiplied by a constant A which depends on p but not on g or the function $\epsilon(s)$ (see [1], Chapter II, Theorem 1). More generally, the same holds for

$$e^{irs}\int_{|s-t|>\epsilon(s)} e^{-irt} g(t)/(s-t)\,dt,$$

with the same constant as before. Thus if $\mu(r)$ is a function of bounded variation in $-\infty < r < \infty$ from Minkowski's integral inequality it follows that the L^p norm of the function of s given by

$$\int_{-\infty}^{+\infty} e^{irs}\Big[\int_{|s-t|>\epsilon(s)} e^{-irt} g(t)/(s-t)\,dt\Big]\,d\mu(r)$$

is not larger than the L^p norm of g multiplied by the constant A above and by the total variation of μ. Now interchanging the order of integration in the expression above (which we may) and observing that the function $\epsilon(s)$ is positive and measurable but otherwise arbitrary we conclude that if

3.1 $$\bar{g}(s) = \sup_{\epsilon} \Big|\int_{|s-t|>\epsilon} \{\psi(s-t)/(s-t)\}\,g(t)\,dt\Big|,$$

where $\psi(s) = \int_{-\infty}^{+\infty} e^{isr}\,d\mu(r)$, then $\|\bar{g}\|_p \leqq AV(\mu)\,\|g\|_p$, $V(\mu)$ being the total variation of μ and A being a constant which only depends on p.

Let now $f(x)$ be a given function of L^p, $1 < p < \infty$, in E_n, y' a unit vector, and define

3.2 $$\tilde{f}_\epsilon(x, y') = \int_{|t|>\epsilon} t^{-1} f(x - ty')\,\psi(t)\,dt,$$

3.3 $$\bar{f}(x, y') = \sup_{\epsilon} |\tilde{f}_\epsilon(x, y')|.$$

Clearly $\bar{f}_\epsilon$ exists for almost all (x, y') and is a measurable function of (x, y'). Furthermore, for almost all (x, y') it is a continuous function of ϵ, so that if we restrict ϵ to rational values in 3.3 we obtain the same value for $\bar{f}$ almost everywhere in (x, y'), which shows that $\bar{f}$ is also measurable.

Now it is readily seen that $\bar{f}(x, y')$ restricted to any straight line parallel to y' is precisely the integral in 3.1 of the function $f(x)$ restricted to the same line. Consequently

$$\int_{-\infty}^{+\infty} \bar{f}(x - ty', y')^p \, dt \leqq A^p V(\mu)^p \int_{-\infty}^{+\infty} |f(x - ty')|^p \, dt,$$

and integrating this inequality over the space of lines parallel to y' we obtain

3.4 $$\int \bar{f}(x, y')^p \, dx \leqq A^p V(\mu)^p \int |f(x)|^p \, dx.$$

Define now

3.5 $$f^\dagger(x) = \tfrac{1}{2} \int_\Sigma \bar{f}(x, y') F(y') \, dy',$$

3.6 $$\tilde{f}_\epsilon(x) = \tfrac{1}{2} \int_\Sigma \tilde{f}_\epsilon(x, y') N(x, y') \, dy',$$

where F and N are the functions introduced in Theorem 3. On account of 3.3 it follows that $|\tilde{f}_\epsilon(x)| \leqq f^\dagger(x)$, and Minkowski's integral inequality applied to 3.5, and 3.4 gives

3.7 $$\| f^\dagger \|_p \leqq \tfrac{1}{2} A V(\mu) \int_\Sigma F(y') \, dy' \, \| f \|_p.$$

But the function $\tilde{f}_\epsilon(x)$ defined in 3.6 coincides with the integral 1.1 as specified in Theorem 3. To see this one merely has to substitute $\tilde{f}_\epsilon(x, y')$ for its value in 3.6 and observe that one obtains 1.1 in polar coordinates with origin at x. Interchanging the order of integration is permissible wherever 1.1 is absolutely convergent, that is, almost everywhere. Thus we have proved that under the assumptions of Theorem 3, $\bar{f}(x) = \sup_\epsilon |\tilde{f}_\epsilon(x)|$ belongs to L^p, and that $\| \bar{f} \|_p \leqq A \| f \|_p$. A more explicit estimate of the constant involved appears in the right-hand side of 3.7, where A depends only on p.

We now prove that the same holds under the assumptions of Theorem 4.

We redefine $f^\dagger(x)$ and $\tilde{f}_\epsilon(x)$ by means of the formulas

$$f^\dagger(x) = \tfrac{1}{2} \int_\Sigma \bar{f}(x, y') | N(x, y') | \, dy', \qquad \tilde{f}_\epsilon(x) = \tfrac{1}{2} \int_\Sigma \tilde{f}_\epsilon(x, y') N(x, y') \, dy'.$$

First we observe that the $\tilde{f}_\epsilon(x)$ just introduced coincides with the $\tilde{f}_\epsilon(x)$

in 1.1. For the last integral above is nothing but 1.1 expressed in polar coordinates with origin at x. On account of 3.3 it follows again that $|\tilde{f}_\epsilon(x)| \leqq f^\dagger(x)$, and Hölder's inequality and Fubini's theorem yield

$$\int f^\dagger(x)^p\, dx = 2^{-p} \int [\int_\Sigma \bar{f}(x,y') \,|\, N(x,y') \,|\, dy']^p\, dx$$

$$\leqq 2^{-p} \int [\int_\Sigma \bar{f}(x,y')^p\, dy'][\int_\Sigma |\, N(x,y') \,|^{p'}\, dy']^{p/p'}\, dx$$

$$\leqq 2^{-p} \int_\Sigma dy' [\int \bar{f}(x,y')^p [\int_\Sigma |\, N(x,y') \,|^{p'}\, dy']^{p/p'}\, dx],$$

where $p' = p/(p-1)$. Now $p' \leqq q$, so that condition i) in Theorem 4 implies that the innermost integral in the last expression above is bounded. Hence if $B^{p'}$ is an upper bound for this integral and ω is the "area" of the unit sphere in E_n, 3.4 yields $\| f^\dagger \|_p \leqq \frac{1}{2} A V(\mu) \omega^{1/p} B \| f \|_p$. Thus we find again that $\| \tilde{f}(x) \|_p \leqq A \| f \|_p$.

Now we can prove that $\tilde{f}_\epsilon(x)$ converges in the mean and pointwise almost everywhere. The argument clearly covers both Theorem 3 and Theorem 4.

Let $\rho(t)$ be an even and continuously differentiable function equal to 1 for $t = 0$ and vanishing outside the interval $(-1, 1)$. It is readily seen that the Fourier transform of $\psi(t)\rho(t)$ is bounded and integrable. Consider now the function equal to t^{-1} for $|\, t \,| > \epsilon$ and zero otherwise. An easy computation shows that its Fourier transform is bounded uniformly in ϵ and converges pointwise as $\epsilon \to 0$. Consequently it follows from Parseval's formula that

3.8 $$\int_{|t|>\epsilon} \psi(t)\rho(t)/t\, dt$$

converges as $\epsilon \to 0$. Thus under the hypotheses of either Theorem 3 or Theorem 4, the integral

$$\int_{|x-y|>\epsilon} K(x,y)\rho(|\, x - y \,|)\, dy$$

converges as $\epsilon \to 0$. To see this one merely has to compute this integral in polar coordinates and use the fact pointed out above that 3.8 converges.

Let now $f(x)$ be continuously differentiable and vanish outside a bounded set. Then

$$\tilde{f}_\epsilon(x) = \int_{|x-y|>\epsilon} K(x,y) f(y)\, dy = \int_{|x-y|>\epsilon} K(x,y)\, [f(y) - f(x)\rho(|\, x - y \,|)]\, dy$$

$$+ f(x) \int_{|x-y|>\epsilon} K(x,y)\rho(|\, x - y \,|)\, dy.$$

The integrand in the first integral on the right is absolutely integrable over E_n, and the second integral converges as $\epsilon \to 0$. Consequently $\tilde{f}_\epsilon(x)$ converges.

In the general case, given $f(x)$ in L^p and $\delta > 0$ there exists a continuously differentiable g vanishing outside a bounded set such that $f = g + h$ and $\| h \|_p < \delta$. Since

$$\tilde{f}_\epsilon(x) = \tilde{g}_\epsilon(x) + \tilde{h}_\epsilon(x), \quad | \tilde{h}_\epsilon(x) | \leqq \bar{h}(x)$$

and $\| \bar{h} \|_p \leqq A \| h \|_p < A\delta$, and since $\tilde{g}_\epsilon(x)$ converges everywhere,

$$\overline{\lim}\, \tilde{f}_\epsilon(x) - \underline{\lim}\, \tilde{f}_\epsilon(x) \leqq 2\bar{h}(x),$$

and this implies that $\tilde{f}_\epsilon(x)$ converges almost everywhere because $\bar{h}(x)$ has arbitrarily small L^p norm. Finally, since $\tilde{f}_\epsilon(x) \to \tilde{f}(x)$ almost everywhere and $| \tilde{f}_\epsilon(x) | \leqq \bar{f}(x)$, the theorem on dominated convergence yields $\| \tilde{f}_\epsilon - \tilde{f} \|_p \to 0$ as $\epsilon \to 0$.

Theorems 3 and 4 are thus established.

4. The proof of Theorems 6 and 7 is based on the same technique we used in the preceding section.

Let $f(t) \geqq 0$ be a function defined in $-\infty < t < \infty$, and $\phi(t) = \psi(t) t^{n-1}$, where $\psi(t)$ is the function introduced in Theorem 6. Then $\phi(t)$ is integrable in $(0, \infty)$. Set

$$f^*(s) = \sup_\epsilon \epsilon^{-1} \int_0^\infty f(s+t)\phi(t\epsilon^{-1})\,dt, \qquad \epsilon > 0,$$

$$F_s(t) = t^{-1} \int_0^t f(s+t)\,dt, \qquad t > 0,$$

and $G(s) = \sup_t F_s(t)$. Then integration by parts gives

$$\epsilon^{-1} \int_0^\infty f(s+t)\phi(t\epsilon^{-1})dt = -\epsilon^{-1} \int_0^\infty tF_s(t)d\phi(t\epsilon^{-1}) \leqq -\epsilon^{-1} G(s) \int_0^\infty t d\phi(t\epsilon^{-1})$$

$$= G(s) \int_0^\infty \phi(t)dt,$$

and consequently $f^*(s) \leqq G(s) \int_0^\infty \phi(t)\,dt$.

Now, a theorem of Hardy and Littlewood (see [3], p. 244), asserts that if $f \varepsilon L^p$, $1 < p < \infty$, then $G \varepsilon L^p$ and $\| G \|_p \leqq A \| f \|_p$, where A depends on p only. Therefore $\| f^* \|_p \leqq A \| f \|_p$, A now depending on p and ϕ.

Let now $f(x) \geqq 0$ be a function from L^p, $1 < p < \infty$ in E_n and y' a unit vector. Define

$$f^*(x, y') = \sup_\epsilon \epsilon^{-1} \int_0^\infty f(x + ty')\phi(t\epsilon^{-1})\,dt.$$

Then, as we have shown above,

$$\int_{-\infty}^{+\infty} f^*(x+ty',y')^p\,dt \leqq A \int_{-\infty}^{+\infty} f(x+ty')^p\,dt,$$

where A depends only on p and ϕ. Integrating over the space of lines parallel to y' we obtain

$$\int f^*(x,y')^p\,dx \leqq A \int f(x)^p\,dx. \tag{4.1}$$

Under the assumptions of Theorem 6 we have

$$f^*(x) = \sup_\epsilon \left| \int K_\epsilon(x,y)f(y)dy \right| \leqq \sup_\epsilon \epsilon^{-n} \int N(x-y)\psi(|x-y|\epsilon^{-1})f(y)dy$$

$$= \sup_\epsilon \int_\Sigma N(y')[\epsilon^{-1}\int_0^\infty f(x+ty')\phi(t\epsilon^{-1})dt]dy' \leqq \int_\Sigma N(y')f^*(x,y')dy',$$

and an application of Minkowski's integral inequality and 4.1 yield

$$\| f^* \|_p \leqq A \int_\Sigma N(y')dy' \| f \|_p,$$

where A depends on p and ψ only.

On the other hand, under the assumptions of Theorem 7 we have

$$f^*(x) = \sup_\epsilon \left| \int K_\epsilon(x,y)f(y)dy \right|$$

$$\leqq \sup_\epsilon \epsilon^{-n} \int |N(x,x-y)|\,\psi(|x-y|\epsilon^{-1})f(y)dy$$

$$= \sup_\epsilon \int_\Sigma |N(x,y')|[\epsilon^{-1}\int_0^\infty f(x+ty')\phi(t\epsilon^{-1})dt]dy'$$

$$\leqq \int_\Sigma |N(x,y')|\,f^*(x,y')dy'.$$

Hence

$$\int f^*(x)^p\,dx \leqq \int [\int_\Sigma |N(x,y')|\,f^*(x,y')dy']^p\,dx$$

$$\leqq \int [\int_\Sigma f^*(x,y')^p\,dy'][\int_\Sigma |N(x,y')^{p'}dy']^{p/p'}\,dx,$$

where $p' = p/(p-1)$. But $p' \leqq q$ so that the integral of $|N(x,y')|^{p'}$ is bounded. If $B^{p'}$ is a bound for this integral and ω is the area of the unit sphere in E_n, interchanging the order of integration and applying 4.1 we obtain finally $\| f^* \|_p \leqq A\omega^{1/p}B \| f \|_p$, where A only depends on p and ψ.

The proof of Theorems 6 and 7 is thus completed.

Remark. The methods used so far still yield results under slightly less restrictive assumptions about the type of integrability of $f(x)$. For instance if f vanishes outside a bounded set and $|f|\log^+|f|$ is integrable, one can still prove that under the assumptions of either Theorem 3 or Theorem 4, $\tilde{f}_\epsilon(x)$ converges in the mean order 1 on any bounded set.

This result is needed in the next section but only in the special case when $K(x,y)$ is the kernel of M. Riesz (see below), and in this form it is also contained in Theorem 7, Chapter 1 of [1]. We may thus safely omit further details.

5. The proof of Theorems 1 and 2 in their full generality is more complicated. In special cases they are contained in Theorems 3 and 4 respectively. In fact, if in Theorem 1 the function $N(x)$ is such that $N(x)=-N(-x)$, then Theorem 1 reduces to Theorem 3 with $N(x,y)=N(y)$ and $\psi(t)=1$. Similarly, if $N(x,y)$ in Theorem 2 is such that $N(x,y)=-N(x,-y)$, then Theorem 2 reduces to Theorem 4 with the same N and $\psi(t)=1$.

Since it is always possible to decompose the functions $N(x)$ and $N(x,y)$ in the sum of two,

$$N(x)=N_1(x)+N_2(x), \qquad N(x,y)=N_1(x,y)+N_2(x,y),$$

where

$$N_1(x)=N_1(-x), \qquad N_2(x)=-N_2(-x),$$

and

$$N_1(x,y)=N_1(x,-y), \qquad N_2(x,y)=-N_2(x,-y),$$

we need only treat the cases $N(x)=N(-x)$, $N(x,y)=N(x,-y)$ and for this purpose we shall use the device of representing f as a singular integral with the kernel of M. Riesz. Our original integral 1.1 will then appear as an iterated integral to which we shall be able to apply the preceding results.

In what follows we shall use vector valued functions but we shall introduce no special notation for them. When talking about the L^p norm of a vector valued function we shall be meaning the L^p norm of its absolute value. The inner product of two vectors will be denoted by their symbols with a dot in-between.

The kernel R of M. Riesz is vector valued and odd

$$R(x)=\pi^{-\frac{1}{2}(n+1)}\Gamma(\tfrac{1}{2}n+\tfrac{1}{2})x\,|\,x\,|^{-n-1}.$$

If

$$g_\epsilon(x)=-\int_{|x-y|>\epsilon}R(x-y)f(y)dy,$$

and $f \varepsilon L^p$, $1 < p < \infty$, then $g_\epsilon(x)$ is a vector valued function which, as $\epsilon \to 0$, converges in the mean of order p to a function $g(x)$. This follows from Theorem 1 by applying it to each component. On the other hand, if

$$f_\epsilon(x) = \int_{|x-y|>\epsilon} R(x-y) \cdot g(y) dy,$$

where the integrand is the inner product of the vectors displayed, then $f_\epsilon(x)$ converges likewise to $f(x)$. That it converges to a function h is again a consequence of Theorem 1. That $h = f$ can be verified for f bounded and vanishing outside a bounded set by taking Fourier transforms (see [2]), whence the general case follows from the continuity in L^p of the linear operation taking f into h.

Let $\phi(t)$ be a continuously differentiable function of the real variable t, $t \geqq 0$, equal to zero in $(0, \frac{1}{4})$ and to 1 in $(\frac{3}{4}, \infty)$, and $F(x)$ a homogeneous function of degree $-n$, such that $F(x) = F(-x)$ and that $|F| \log^+ |F|$ is integrable on the sphere $|x| = 1$. Suppose in addition that the integral of $F(x)$ over $|x| = 1$ is zero and consider

5.1 $$\int_{|x-y|>\epsilon} R(x-y) F(y) dy,$$

5.2 $$\int_{|x-y|>\epsilon} R(x-y) F(y) \phi(|y|) dy.$$

Since $|R(x)| \leqq A |x|^{-n}$ the second integral is absolutely convergent. The first has a singularity at $y = 0$, but, owing to the fact that $R(x)$ is continuously differentiable if $x \neq 0$, it can be given a natural meaning if $|x| > \epsilon$ by integrating outside a small sphere with center at $y = 0$ and taking the limit of the value obtained as the radius of the small sphere tends to zero.

The properties of the integrals above which we need are summarized in the following

LEMMA. *Under the preceding assumptions, as* $\epsilon \to 0$, 5.2 *converges in the mean of order* 1 *on any compact set, and* 5.1 *converges on any compact set not containing the point* $x = 0$. *The corresponding limits,* $F_2(x)$ *and* $F_1(x)$, *are odd functions, i.e.* $F_1(x) = -F_1(-x)$, $F_2(x) = -F_2(-x)$. *The function* $F_1(x)$ *is homogeneous of degree* $-n$, *and, for* $|x| \geqq 1$,

5.3 $$|F_1(x) - F_2(x)| \leqq A \int_\Sigma |F(y')| dy' |x|^{-n-1}.$$

There exists a homogeneous function $G(x)$ *of degree zero such that for* $|x| \leqq 1$,

5.4 $$|F_2(x)| \leqq G(x),$$

and

5.5 $$\int_\Sigma G(y')dy' < \infty.$$

If for some q, $1 < q < \infty$, $\int_\Sigma |F(y')|^q dy' < \infty$, *then the inequalities*

5.6 $$\int_\Sigma |F_1(y')|^q dy' \leqq A \int_\Sigma |F(y')|^q dy',$$

5.7 $$\int_\Sigma G(y')^q dy' \leqq A \int_\Sigma |F(y')|^q dy'$$

hold, with the constants A *depending on* q *but not on* F, *and the integral in* 5.2 *converges to its limit in the mean of order* q.

That the functions $F_1(x)$ and $F_2(x)$, if existent, are odd is clear. That $F_1(x)$ is homogeneous of degree $-n$ is also clear. To see that 5.1 converges in the mean between two spheres of radii $\rho_1 < \rho_2$ we observe that the contributions to the integral from the sphere $|y| \leqq \frac{1}{2}\rho_1$ and from the exterior of the sphere $|y| = 2\rho_2$ is bounded, and to the integral extended over $\frac{1}{2}\rho_1 \leqq |y| \leqq 2\rho_2$ we may apply the remark of Section 4 and obtain immediately the desired result. On the other hand, an application of Theorem 4 to each component of the vector valued integral 5.1 gives that the integral of $|F_1(x)|^q$ extended to the region between two spheres with center at $x = 0$ is dominated by the q-th power of the right hand of 5.6. Since $F_1(x)$ is homogeneous, 5.6 follows. A similar argument applies to the integral 5.2, except that in this case it will not be necessary to exclude a neighborhood of $x = 0$.

For the difference between $F_1(x)$ and $F_2(x)$ we get the following estimates

$$\begin{aligned} |F_2(x) - F_1(x)| &\leqq \left| \int R(x-y)F(y)[\phi(|y|) - 1]dy \right| \\ &= \left| \int [R(x-y) - R(x)]F(y)[\phi(|y|) - 1]dy \right| \\ &\leqq \int_{|y|\leqq 3/4} |F(y)|\,|R(x-y) - R(x)|\,dy. \end{aligned}$$

Now, it is readily seen that for $|x| \geqq 1$ and $|y| \leqq \frac{3}{4}$ we have the inequality

$$|R(x-y) - R(x)| \leqq A\,|x|^{-n-1}\,|y|.$$

Substituting this in the preceding integral we obtain 5.3.

In order to prove 5.4, 5.5 and 5.7 we proceed as follows. First we observe that, owing to the fact that $F(y)\phi(|y|)$ vanishes in $|y| \leqq \frac{1}{4}$, $F_2(x)$

is continuous and bounded in $|x| \leqq \frac{1}{8}$, and that $A \int_\Sigma |F(y')|\, dy'$ with an appropriate constant A independent of F is an upper bound for $|F_2(x)|$ in this particular domain $|x| \leqq \frac{1}{8}$.

In $\frac{1}{8} \leqq |x| \leqq 1$ we have

$$\begin{aligned} |F_2(x)| &\leqq \phi(|x|)|F_1(x)| + |F_2(x) - \phi(|x|)F_1(x)| \\ &\leqq |F_1(x)| + \left| \int R(x-y)[\phi(|y|)F(y) - \phi(|x|)F(y)]dy \right| \\ &= |F_1(x)| + \left| \int [R(x-y) - \chi(|y|)R(x)]\, [\phi(|y|) - \phi(|x|)]F(y)dy \right|, \end{aligned}$$

where χ is the characteristic function of the interval $(0,1)$.

Now, one verifies easily that if $\frac{1}{8} \leqq |x| \leqq 1$, then

$$|R(x-y) - \chi(|y|)R(x)| \leqq A\, |y|^{\frac{1}{2}} \cdot |x-y|^{-n}.$$

On the other hand, since $\phi(t)$ has a bounded derivative,

$$|\phi(|x|) - \phi(|y|)| \leqq A\, \big|\, |x| - |y|\, \big| \leqq A\, |x-y|.$$

Thus in the last integral, substituting this we get

$$|F_2(x)| \leqq |F_1(x)| + A \int |y|^{\frac{1}{2}} |F(y)|\, |x-y|^{-(n-1)} dy,$$

and since $|x| \geqq \frac{1}{8}$, it follows that

$$|F_2(x)| \leqq A[\, |x|^n |F_1(x)| + |x|^{n-\frac{3}{2}} \int |y|^{\frac{1}{2}} |F(y)|\, |x-y|^{-(n-1)} dy].$$

The integral on the right represents a homogeneous function of degree $-n + \frac{3}{2}$, and $F_1(x)$ is homogeneous of degree $-n$. Hence the right-hand side of the inequality is a homogeneous function of degree zero. It remains only to show 5.5 and 5.7. The contribution of the term $|x|^n |F_1(x)|$ can be estimated either by the fact that $F_1(x)$ is integrable between two concentric spheres with center at the origin, or by 5.6. The contribution of the other term is estimated by splitting the integral in two, one extended over the set $\frac{1}{16} \leqq |y| \leqq 2$, the other over the complement of this set. The second integral is readily seen to be bounded by $A \int_\Sigma |F(y')|\, dy'$ or $A[\int_\Sigma |F(y')|^q dy']^{1/q}$ and an application of the theorem of Young (see [3], page 71, where the theorem is stated and proved in a special case; but the proof extends obviously to the most general situation) to the first shows that it represents a function of the same class as $F(x)$ in $\frac{1}{16} \leqq |x| \leqq 2$. But the function under consideration is homogeneous of degree zero and, consequently, an estimate for

its norm in $\frac{1}{2} \leqq |x| \leqq \frac{3}{2}$ gives an estimate for its norm on the sphere $|x| = 1$. Collecting results, 5.5 and 5.7 follow and the proof of our lemma is complete.

Let now $N(x)$ have the property that $N(x) = N(-x)$ and satisfy the conditions of Theorem 1. Then, by the lemma above,

$$N_1(x) = \lim_{\epsilon \to 0} \int_{|y-x|>\epsilon} R(x-y) N(y)\, dy$$

also satisfies the conditions of Theorem 1 and $N_1(x) = -N_1(-x)$. Furthermore, if

$$N_2(x) = \lim_{\epsilon \to 0} \int_{|x-y|>\epsilon} R(x-y) N(y) \phi(|y|)\, dy,$$

then, for $|x| \geqq 1$,

5.8 $$|N_2(x) - N_1(x)| \leqq A\, |x|^{-n-1},$$

and for $|x| \leqq 1$,

5.9 $$|N_2(x)| \leqq G(x),$$

where $G(x)$ is a homogeneous function of degree zero integrable over the sphere $|x| = 1$.

Consider now the vector valued function $g(x)$ and

$$f(x) = \lim_{\epsilon \to 0} \int_{|x-y|>\epsilon} R(x-y) \cdot g(y)\, dy.$$

As we already know, if $|g(x)| \varepsilon L^p$, $1 < p < \infty$, then $f(x) \varepsilon L^p$ and $\| f \|_p \leqq A \| g \|_p$. Furthermore the integral above converges in the mean of order p and every function $f(x)$ in L^p can be thus represented.

We shall prove the following identity:

5.10 $$\int N(x-y) \phi(|x-y|\, \epsilon^{-1}) f(y)\, dy = \epsilon^{-n} \int N_2(\epsilon^{-1}(x-y)) \cdot g(y)\, dy.$$

If g is continuously differentiable and vanishes outside a bounded set, then, on account of absolute integrability,

5.11 $$\begin{aligned} &\int N(x-y) \phi(|x-y|\, \epsilon^{-1})\, dy \int_{|y-z|>\delta} R(y-z) \cdot g(z)\, dz \\ &= \int \Big[\int_{|y-z|>\delta} N(x-y) \phi(|x-y|\, \epsilon^{-1}) R(y-z)\, dy \Big] \cdot g(z)\, dz. \end{aligned}$$

Now, as $\delta \to 0$, by the lemma above and by changing variables, the inner integral on the right is seen to converge to $\epsilon^{-n} N_2((x-y)\epsilon^{-1})$ in the mean

of order 1 on any compact set. Therefore the right-hand side above converges to

$$\epsilon^{-n}\int N_2((x-y)\epsilon^{-1})\cdot g(y)\,dy.$$

On the other hand,

$$\int_{|y-z|>\delta} R(y-z)\cdot g(z)\,dz=\int_{|y-z|>\delta} R(y-z)\cdot[g(z)-g(y)]\,dz,$$

and on account of the continuous differentiability of g, the right-hand side is readily seen to converge uniformly as $\delta\to 0$ and to be independent of δ for $\delta<1$ and $|y|$ sufficiently large. Therefore, the left-hand side of 5.11 is seen to converge to

$$\int N(x-y)\phi(|x-y|\epsilon^{-1})f(y)\,dy.$$

Thus 5.10 is proved for g continuously differentiable and vanishing outside a bounded set. In the general case, given $g\,\varepsilon\,L^p$, we take a sequence of continuously differentiable functions g_k, each vanishing outside a compact set and such that $\| g_k-g\|_p\to 0$ and $\sum_1^\infty \| g_{k+1}-g_k\|_p<\infty$. If

$$f_k(x)=\lim_{\epsilon\to 0}\int_{|x-y|>\epsilon} R(x-y)\cdot g_k(y)\,dy,$$

then $\| f_k-f\|_p\to 0$ and $\sum_1^\infty \| f_{k+1}-f_k\|_p<\infty$. From the finiteness of the series $\sum\| g_{k+1}-g_k\|_p$ and $\sum\| f_{k+1}-f_k\|_p$ it follows that the functions $\bar g=|g_1|+\sum|g_{k+1}-g_k|$ and $\bar f=|f_1|+\sum|f_{k+1}-f_k|$ are finite almost everywhere and belong to L^p. Thus the sequences

$$g_k(x)=g_1(x)+\sum_1^{k-1}[q_{j+1}(x)-g_j(x)] \text{ and } f_k(x)$$

converge almost everywhere and are dominated in absolute value by $\bar g$ and $\bar f$ respectively. Now the considerations of Section 2 show that the integral

$$\int |N(x-y)|\,\phi(|x-y|\epsilon^{-1})\bar f(y)\,dy\leqq\int_{|x-y|>\epsilon/4}|N(x-y)|\,\bar f(y)\,dy$$

is finite for almost all x. On the other hand, on account of 5.8 we have

$$\int |N_2((x-y)\epsilon^{-1})|\,\bar g(y)\,dy\leqq\int_{|x-y|<\epsilon}|N_2((x-y)\epsilon^{-1})|\,\bar g(y)\,dy$$

$$+\int_{|x-y|>\epsilon}|N_1((x-y)\epsilon^{-1})|\,\bar g(y)\,dy+A\int_{|x-y|>\epsilon}|x-y|^{-n-1}\,\bar g(y)\,dy.$$

The first and last integrals on the right are absolutely convergent for almost all x owing to the absolute integrability of $N_2(x)$ in $|x| \leqq 1$, and the remaining integral, by the considerations of Section 2, is also finite for almost all x. Hence both

$$\int |N(x-y)|\phi(|x-y|\epsilon^{-1})\bar{f}(y)dy$$

and

$$\int |N_2((x-y)\epsilon^{-1})|\bar{g}(y)dy$$

are finite for almost all x. Consequently we can pass to the limit in

$$\int N(x-y)\phi(|x-y|\epsilon^{-1})f_k(y)dy = \epsilon^{-n}\int N_2((x-y)\epsilon^{-1})\cdot g_k(y)dy,$$

and we find that in the general case 5.10 holds for almost all x.

Now the proof of Theorem 1 is nearly completed. We have

$$\begin{aligned}\tilde{f}_\epsilon(x) &= \int_{|x-y|>\epsilon} N(x-y)f(y)dy = \int N(x-y)\phi(|x-y|\epsilon^{-1})f(y)dy \\ &\quad - \int_{|x-y|<\epsilon} N(x-y)\phi(|x-y|\epsilon^{-1})f(y)dy \\ &= \epsilon^{-n}\int N_2((x-y)\epsilon^{-1})\cdot g(y)dy - \int_{|x-y|<\epsilon} N(x-y)\phi(|x-y|)\epsilon^{-1})f(y)dy,\end{aligned}$$

and on account of 5.8 and 5.9 we find that

$$\begin{aligned}|\tilde{f}_\epsilon(x)| \leqq \left|\int_{|x-y|>\epsilon} N_1(x-y)\cdot g(y)dy\right| &+ \epsilon^{-n}\int_{|x-y|<\epsilon} G((x-y)/|x-y|)|g(y)|\,dy \\ &+ A\epsilon^{-n}\int_{|x-y|>\epsilon} |(x-y)\epsilon^{-1}|^{-n-1}|g(y)|\,dy \\ &+ A\epsilon^{-n}\int |N((x-y)/|x-y|)||f(y)|\,dy,\end{aligned}$$

whence Theorem 1 applied to the first term on the right, and Theorem 6 applied to the remaining ones yield

$$\|\tilde{f}\|_p = \|\sup_\epsilon |\tilde{f}_\epsilon|\|_p \leqq A\|g\|_p + A\|f\|_p \leqq A\|f\|_p.$$

From this convergence in the mean and almost everywhere of $\tilde{f}_\epsilon(x)$ follows as in the proof of Theorems 3 and 4. Theorem 1 is thus proved.

The proof of Theorem 2 proceeds along similar lines but the differences justify its presentation. Let $K(x,y)$ be as specified in Theorem 2 with the additional property that $K(x,y) = K(x,-y)$ and define

$$K_1(x,y)=\lim_{\epsilon\to 0}\int_{|y-z|>\epsilon} K(x,z)R(y-z)\,dz,$$

$$K_2(x,y)=\lim_{\epsilon\to 0}\int_{|y-z|>\epsilon} K(x,z)\phi(|z|)R(y-z)\,dz.$$

Both K_1 and K_2 are odd functions in y, and K_1 satisfies the conditions of Theorem 2. Furthermore, for $|y|\geqq 1$,

5.12 $$|K_2(x,y)-K_1(x,y)|\leqq A\,|y|^{-n-1},$$

with A independent of x, and for $|y|\leqq 1$

5.13 $$|K_2(x,y)|\leqq G(x,y),$$

where G is homogeneous of degree zero in y and such that

5.14 $$\int_{\Sigma} G(x,y')^q\,dy'$$

is bounded. All this follows from our lemma. Let now $g(x)$ be a vector valued function in L^p, $p\geqq q/(q-1)$ and set

$$f(x)=\lim_{\epsilon\to 0}\int_{|x-y|>\epsilon} R(x-y)\cdot g(y)\,dy.$$

We shall prove the identity

5.15 $$\int K(x,x-y)\phi(|x-y|\,\epsilon^{-1})f(y)\,dy = \epsilon^{-n}\int K_2(x,(x-y)\epsilon^{-1})\cdot g(y)\,dy.$$

If $g(y)$ is continuously differentiable and vanishes outside a bounded set this identity is proved the same way we proved 5.10. In the general case, given $g\,\varepsilon\,L^p$ we take a sequence of continuously differentiable functions g_n, each vanishing outside a bounded set, and tending to g in the mean of order p. Since both $K(x,x-y)\phi(|x-y|\,\epsilon^{-1})$ and $K_2(x,(x-y)\epsilon^{-1})$ as functions of y are of integrable power $p/(p-1)$, and since the functions f_n corresponding to the g_n converge to f in the mean of order p, a passage to the limit under the integral sign yields 5.15 in its full generality.

Now we have

$$\begin{aligned}\tilde f_\epsilon(x)&=\int_{|x-y|>\epsilon} K(x,x-y)f(y)\,dy=\int K(x,x-y)\phi(|x-y|\,\epsilon^{-1})f(y)\,dy\\&\quad-\int_{|x-y|<\epsilon} K(x,x-y)\phi(|x-y|\,\epsilon^{-1})f(y)\,dy\\&=\epsilon^{-n}\int K_2(x,(x-y)\epsilon^{-1})\cdot g(y)\,dy-\int_{|x-y|<\epsilon} K(x,x-y)\phi(|x-y|\,\epsilon^{-1})f(y)\,dy,\end{aligned}$$

and on account of 5.12 and 5.13 we find that

$$|\tilde{f}_\epsilon(x)| \leqq \left| \int_{|x-y|>\epsilon} K_1(x, x-y) \cdot g(y) dy \right| + \epsilon^{-n} \int_{|x-y|<\epsilon} G(x, (x-y)/|x-y|) |g(y)| dy$$
$$+ A\epsilon^{-n} \int_{|x-y|>\epsilon} |(x-y)\epsilon^{-1}|^{-n-1} |g(y)| dy + A\epsilon^{-n} \int_{|x-y|<\epsilon} |K(x, (x-y)/|x-y|)| |f(y)| dy,$$

whence Theorem 2 applied to the first term on the right and Theorem 7 applied to the remaining ones yield

$$\| \tilde{f} \|_p = \| \sup_\epsilon |\tilde{f}_\epsilon(x)| \|_p \leqq A(\| g \|_p + \| f \|_p) \leqq A \| f \|_p.$$

The argument can now be completed as before.

We close this section by showing that Theorem 2 ceases to hold for functions in L^p, $p < q/(q-1)$.

Let $p < q/(q-1)$ and take an integer n and $\alpha > 0$ so that

$$1/\alpha \leqq p/(p-1) - q, \qquad n/\alpha = p/(p-1).$$

Define $f(x)$ in E_n as follows: $f(x) = 1$ for $|x| \leqq 1$ and $f(x) = 0$ otherwise. Set

$$K(x,y) = |x|^\alpha |y|^{-n} \text{ for } |x| \geqq 1 \text{ and } |x' + y'| \leqq 1/|x|,$$
$$K(x,y) = -|x|^\alpha |y|^{-n} \text{ for } |x| \geqq 1 \text{ and } |x' - y'| \leqq 1/|x|$$

and $K(x,y) = 0$ otherwise. Then one sees readily that

$$\int_\Sigma |K(x,y')|^q dy' \leqq A |x|^{\alpha q - n + 1} \qquad |x| \geqq 1$$

But

$$\alpha q - n + 1 = \alpha q - \alpha p/(p-1) + 1 = \alpha[q - p/(p-1)] + 1 \leqq 0,$$

and consequently the integral above is bounded. On the other hand,

$$|\tilde{f}(x)|^p \geqq A |x|^{(\alpha - n)p} \text{ as } |x| \to \infty,$$

and

$$(n-\alpha)p = \alpha p/(p-1) = n,$$

so that $\tilde{f} \notin L^p$.

Remark. In order to simplify our presentation as far as possible we have omitted to give explicit estimates for the constant A in the inequalities

$\| \tilde{f} \|_p \leqq A \| f \|_p$. In the paper that follows this, though, we shall need to know more about A. If, in Theorem 1, $N(x)$ is such that

$$\int_\Sigma | N(y') |^q \, dy' < \infty, 1 < q < \infty,$$

then

$$A = A_{pq} [\int_\Sigma | N(y') |^q \, dy']^{1/q},$$

where A_{pq} depends on p and q but *not* on $N(x)$. The reader will have little difficulty in verifying this statement himself, by estimating step by step the constants in the preceding proofs.

6. In this last section we shall prove Theorem 5. We restrict ourselves to the case of three or more variables.

Let $f(x)$ be a function L^p, $1 < p \leqq 2$. Then $f(x)$ has a Fourier transform f given by

6.1 $$\hat{f}(x) = \lim_{r \to \infty} \int_{|y| < r} e^{i(x \cdot y)} f(y) \, dy,$$

the limit being understood as a limit in the mean of order $p/(p-1)$. The spherical means of order $\frac{1}{2}(n-1)$ of the Fourier integral of f are given by

6.2 $$\sigma_r(f, y) = (2\pi)^{-n} \int_{|x| < r} \hat{f}(x) e^{-i(x \cdot y)} (1 - | x |^2 r^{-2})^{\frac{1}{2}(n-1)} \, dx.$$

If we assume that f vanishes outside a compact set we may replace $\hat{f}$ by its value 6.1 and interchange the order of integration obtaining

6.3 $$\sigma_r(f, y) = (2\pi)^{-n} \int [\int_{|x| < r} e^{ix \cdot (z-y)} (1 - | x |^2 \, r^{-2})^{\frac{1}{2}(n-1)} \, dx] f(z) \, dz.$$

If we set

$$F(z) = \int_{|y| \leqq 1} e^{-i(y \cdot z)} (1 - | y |^2)^{\frac{1}{2}(n-1)} \, dy,$$

6.3 becomes

6.4 $$\sigma_r(f, y) = (2\pi)^{-n} r^n \int F[r(y-z)] f(z) \, dz.$$

Now the function $F(z)$ is in L^2 and bounded and consequently it belongs to L^q for every $q \geqq 2$. Therefore 6.4 can be extended to an arbitrary f in L^p by a passage to the limit.

Our next step will be to prove that

$$\| \sigma_r(f, y) \|_p \leqq A \| f \|_p.$$

For this purpose we shall show that

$$F(z) = \psi(|z|)\,|z|^{-n}$$

where $\psi(t)$ is an odd Fourier-Stieltjes integral. We take an orthogonal coordinate system in E_n whose first coordinate axis coincides with z and denote by t the corresponding coordinate. Then

$$|z|^n F(z) = |z|^n \int_{|y|\leqq 1} e^{-i(y\cdot z)}(1-|y|^2)^{\frac{1}{2}(n-1)}\,dy$$

$$= |z|^n \omega_{n-2} \int_{-1}^{+1} e^{i|z|t}\Big[\int_0^{(1-t^2)^{\frac{1}{2}}} [1-(t^2+s^2)]^{\frac{1}{2}(n-1)} s^{n-2}\,ds\Big]dt,$$

where ω_{n-2} denotes the "area" of the $n-1$ dimensional unit sphere.

Now by setting $s^2 = v(1-t^2)$ the inner integral on the right is readily seen to be equal to

$$\tfrac{1}{2}(1-t^2)^{n-1}\int_0^1 (1-v)^{\frac{1}{2}(n-1)} v^{\frac{1}{2}(n-3)}\,dv,$$

and thus

$$|z|^n F(z) = A\,|z|^n \int_{-1}^{+1} e^{i|z|t}(1-t^2)^{n-1}\,dt = \psi(|z|),$$

where

$$\psi(t) = At^n \int_{-1}^{+1} e^{ist}(1-s^2)^{n-1}\,ds.$$

Since n is odd, $\psi(t)$ is odd and an n-fold integration by parts shows that $\psi(t)$ is a Fourier-Stieltjes transform.

Thus 6.4 becomes

$$6.5 \qquad \sigma_r(f,y) = (2\pi)^{-n}\int \psi(r\,|y-z|)\,|y-z|^{-n} f(z)\,dz.$$

Since, for each r, $\psi(rt)$ is the Fourier-Stieltjes transform of a function whose total variation is independent of r, Theorem 3 applied to the preceding integral yields

$$6.6 \qquad \|\sigma_r(f,y)\|_p \leqq A\,\|f\|_p,$$

with A independent of r.

Suppose now that $f(x)$ has continuous derivatives of all orders and vanishes outside a bounded set. Then $\hat{f}$ is absolutely integrable and the inversion theorem for Fourier transforms shows that $\sigma_r(f,y)$ converges to $f(y)$ as $r\to\infty$. Furthermore $\sigma_r(f,y)$ is bounded uniformly in r and an inspection of 6.5 shows that

$$|\sigma_r(f,y)| \leqq A\,|y|^{-n}$$

for $|y|$ sufficiently large, regardless of the value of r. This makes it clear that

6.7 $$\| \sigma_r(f,y) - f(y) \|_p \to 0$$

as $r \to \infty$.

In the general case, given $f \varepsilon L^p$ and $\epsilon > 0$ there exists a function g with continuous derivatives of all orders and vanishing outside a bounded set such that $\| f - g \|_p < \epsilon$. Then

$$\| \sigma_r(f,y) - f(y) \|_p \leqq \| \sigma_r(g,y) - g(y) \|_p + \| \sigma_r(f-g,y) \|_p + \| f - g \|_p,$$

and according to 6.6 and 6.7 as $r \to \infty$ the right-hand side has an upper limit not exceeding $(A+1)\epsilon$. Since ϵ is arbitrary, the proof is complete.

In closing this section we point out that if $\sigma_r(f,y)$ is defined directly by means of 6.4, then the theorem holds for $f \varepsilon L^p$ for any p, $1 < p < \infty$. We also remark that the same method of proof can be used to establish the corresponding result for other appropriate methods of spherical summation.

MASSACHUSETTS INSTITUTE OF TECHNOLOGY AND
THE UNIVERSITY OF CHICAGO.

REFERENCES.

[1] A. P. Calderón and A. Zygmund, "On the existence of certain singular integrals," *Acta Mathematica*, vol. 88 (1952), pp. 85-139.

[2] J. Horvàth, "Sur les fonctions conjuguées à plusieurs variables," Koninkligke Nederlandse Akademie van Wetenschappen. Indagationes Mathematicae ex Actis Quibus Titulis. *Proceedings of the Section of Sciences*, vol. 15, No. 1 (1953), pp. 17-29.

[3] A. Zygmund, *Trigonometrical series*, Warsaw (1935).

ALGEBRAS OF CERTAIN SINGULAR OPERATORS.*

By A. P. CALDERÓN and A. ZYGMUND.[1]

1. In this note we study composition of singular integral operators of a type we have considered in earlier work.

Let $x = (\xi_1, \xi_2, \cdots, \xi_n)$ denote either a point of Euclidean n-space of coordinates $\xi_1, \xi_2, \cdots, \xi_n$, or the vector from $0 = (0, 0, \cdots, 0)$ to $(\xi_1, \xi_2, \cdots, \xi_n)$, and $|x|$ its length, that is $(\xi_1^2 + \cdots + \xi_n^2)^{\frac{1}{2}}$.

If $K(x)$ is a homogeneous function of degree $-n$, i.e. such that

$$K(\lambda x) = \lambda^{-n} K(x)$$

for every x and every $\lambda > 0$, and if

$$\int_\Sigma K(x)\, d\sigma = 0 \quad \text{and} \quad \int_\Sigma |K(x)|^p\, d\sigma < \infty \tag{1.1}$$

for some $p > 1$, the integral being taken over the unit sphere Σ, $|x| = 1$, and $d\sigma$ denoting the elements of "area" of Σ, then for $f \varepsilon L^r$

$$\tilde{f}_\epsilon(x) = \int_{|x-y|>\epsilon} K(x-y) f(y)\, dy \tag{1.2}$$

converges pointwise almost everywhere and in the mean order r as ϵ tends to zero, and the operation of taking f into the limit $\tilde{f}$ of the integral above is continuous in L^r, and

$$\|\tilde{f}\|_r \leqq A_{r,p}\left[\int_\Sigma |K(x)|^p\, d\sigma\right]^{1/p} \|f\|_r, \tag{1.3}$$

where A_{rp} is a constant depending on p and r only (see [3], remark to Section 5).

This result suggests studying composition of operators of the form

$$\mathcal{K}(f) = \alpha f + \tilde{f}, \tag{1.4}$$

where α is a complex constant.

We shall consider

i) the class $\mathfrak{A}$ of all operators with $K(x)$ in C^∞ in $|x| > 0$, that is with $K(x)$ possessing derivatives of all orders if $x \neq 0$;

* Received September 9, 1955.

[1] The research resulting in this paper was supported in part by the office of Scientific Research of the Air Force under contract AF 18(600)-1111.

ii) the class $\mathfrak{A}_p$, $p > 1$, of all operators $\mathcal{K}$ for which

$$\|\mathcal{K}\|_p = |\alpha| + \left[\int_\Sigma |K(x)|^p \, d\sigma\right]^{1/p} < \infty. \tag{1.5}$$

Such classes are obviously closed under addition and multiplication by scalars. To show that they are also closed under composition (operator multiplication) is one of the purposes of this note. Moreover, we intend to prove that $\mathfrak{A}_p$, when endowed with the norm (1.5), is a commutative semi-simple Banach algebra. This will follow from the inequality

$$\|\mathcal{K}\mathcal{H}\|_p \leqq A_p \|\mathcal{K}\|_p \|\mathcal{H}\|_p, \tag{1.6}$$

where A_p is a constant depending on p only.

Since all operators under consideration are continuous in L^2 and commute with translations, if $\mathcal{F}(g)$ denotes the Fourier transform of g, and $f \varepsilon L^2$, we must have

$$\mathcal{F}[\mathcal{K}(f)] = \mathcal{F}(f)\mathcal{F}(\mathcal{K}),[2] \tag{1.7}$$

where $\mathcal{F}(\mathcal{K})$ is a bounded function which we shall call the *Fourier transform of* $\mathcal{K}$. Further, and according to (1.3) and (1.5), and assuming that the constant $A_{r,p}$ in (1.3) is $\geqq 1$,

$$|\mathcal{F}(\mathcal{K})| \leqq A_{2,p} \|\mathcal{K}\|_p. \tag{1.8}$$

As we shall see, $\mathcal{F}(\mathcal{K})$ is actually a homogeneous function of degree zero, continuous in $x \neq 0$. If $\mathcal{K} \varepsilon \mathfrak{A}$, then $\mathcal{F}(\mathcal{K})$ is in addition of class C^∞ in $x \neq 0$. Conversely, every homogeneous function of degree zero possessing derivatives of all orders in $x \neq 0$ is the Fourier transform of an operator in $\mathfrak{A}$.

Finally we shall prove that an operator in $\mathfrak{A}$ or $\mathfrak{A}_p$ has an inverse in the same class if and only if its Fourier transform does not vanish. Since we may identify homogeneous functions of degree zero with their restrictions to the sphere $|x| = 1$, we can translate the last statement into the language of Banach Algebras and assert that the space of maximal ideals of $\mathfrak{A}_p$ is homeomorphic with the sphere $|x| = 1$.

For the convenience of the reader we summarize our results in the following formal statements.

---

[2] In our special case this known general statement also follows from the fact that the Fourier transform of $K_\lambda(x)$, defined to be $K(x)$ for $|x| > \lambda$ and zero otherwise, converges boundedly to a bounded function as $\lambda \to 0$ (see [2], pp. 89-91). For if $f \varepsilon L^2$, the Fourier transform of (1.1) converges in L^2 to the product of a bounded function depending on K only and the Fourier transform of f.

THEOREM 1. *If $\mathfrak{A}$ is the class of all operators defined in* i), *then $\mathfrak{A}$ is closed under addition and operator multiplication. The Fourier transform of an operator in this class is a homogeneous function of degree zero and of C^{∞} in $x \neq 0$, and conversely every such homogeneous function is the Fourier transform of an operator in $\mathfrak{A}$. The Fourier transform of the product of two operators is the product of their Fourier transforms, and consequently an operator in $\mathfrak{A}$ has an inverse in $\mathfrak{A}$ if and only if its Fourier transform does not vanish. If $k(x)$ is the Fourier transform of the kernel $K(x)$, and $\beta(k)$ is the least common upper bound for the absolute value of k and of its derivatives up to order $n+1$ evaluated in $|x| \geqq 1$, then for $|x|=1$ we have*

$$|K(x)| \leqq A\beta(k), \tag{1.9}$$

where A is a constant independent of K.

THEOREM 2. *If $\mathfrak{A}_p$ is the class of all operators defined in* ii) *and endowed with the norm* (1.5), *then $\mathfrak{A}_p$ becomes a semisimple commutative Banach Algebra under operator multplication, and* (1.6) *holds for the norm of the product of two operators. Then Fourier product of an operator in $\mathfrak{A}_p$ is a homogeneous function of degree zero continuous in $|x| \neq 0$, and the Fourier transform of a product is the product of the Fourier transforms of the factors.*

The existence of inverses and the functional calculus of operators in $\mathfrak{A}_p$ is based on the following two theorems.

THEOREM 3. *Let $g(x)$ be a function defined on the sphere Σ ($|x|=1$) which is locally a restriction to Σ of Fourier transforms of operators in $\mathfrak{A}_p$; that is, every $x_0 \in \Sigma$ is contained in a neighborhood where $g(x)$ coincides with the restriction to Σ of the Fourier transform of an operator in $\mathfrak{A}_p$. Then there exists a single operator $\mathfrak{H}$ in $\mathfrak{A}_p$ whose Fourier transform coincides with g at all points of Σ.*

THEOREM 4. *Let $g(x)$ be a function defined on the sphere Σ, which is locally an analytic function of the restriction $h(x)$ to Σ of the Fourier transform of an operator in $\mathfrak{A}_p$; that is, for every $x_0 \in \Sigma$ there exists a power series $\Sigma a_n Z^n$ with positive radius of convergence such that $g(x) = \Sigma a_n[h(x) - h(x_0)]^n$ for x in some neighborhood of x_0. Then $g(x)$ is a restriction to Σ of the Fourier transform of an operator in $\mathfrak{A}_p$.*

COROLLARY. *An operator in $\mathfrak{A}_p$ has an inverse in $\mathfrak{A}_p$ if and only if its Fourier transform does not vanish. The space of maximal ideals of $\mathfrak{A}_p$ is homeomorphic to the sphere Σ ($|x|=1$).*

One thing here must be stressed.

If an operator in $\mathfrak{A}_p$ is thought of as acting in the space L^2 of square integrable functions, then the fact that its Fourier transform does not vanish implies immediately that there is a bounded operator in L^2 which is the inverse of the given one. The fact, however, that this operator is in $\mathfrak{A}_p$ is non-trivial, and this fact is, of course, the essence of the preceding corollary. A similar remark applies to Theorems 2, 3 and 4.

The content of Theorem 4 can be described briefly by saying that an analytic function of the Fourier transform $h(x)$ of an operator in $\mathfrak{A}_p$ is again the Fourier transform of an operator in $\mathfrak{A}_p$. This analytic function need not be single valued, and the values of $h(x)$ might even be allowed to go through branch points of the function, provided that the conditions of Theorem 4 are respected at such points x. For example, if the function $h(x)$ has a continuous square root and coincides locally with Fourier transforms of operators in $\mathfrak{A}_p$ at all points where $h(x)$ vanishes then $h(x)$ has a square root in $\mathfrak{A}_p$.

The preceding theorems apply immediately to systems of singular integral operators in $\mathfrak{A}$ or $\mathfrak{A}_p$. Such systems may be thought of as a convolution of a square singular matrix kernel with a vector function plus a numerical matrix applied to the same function. The condition of invertibility then becomes that the matrix of the corresponding Fourier transforms have a nonvanishing determinant.

2. With things organized as we have them here it will be convenient to study first operators in $\mathfrak{A}$. Once the basic facts about such operators are established and the validity of (1.6) is proved in this special case, everything else will be relatively simple.

The following partly standard notation will be sufficient for our purposes. We shall write $f \cdot g$ for the (absolutely convergent) integral of $f(x)\bar{g}(x)$, $f * g$ for the convolution of f and g, $\mathfrak{F}(f)$ for the Fourier transform of f. We shall also write $g^\lambda(x) = \lambda^n g(\lambda x)$, and denote by $g_\lambda(x)$ the function equal to g if $|x| \geqq \lambda$ and to zero otherwise (the latter notation will apply to kernels only and will not conflict with the notation on the left side of (1.2)).

By Γ we shall denote the class of all functions g of C^∞ such that g and all its derivatives are $O(|x|^{-k})$ as $|x| \to \infty$, for each $k > 0$. The Fourier transform of a function in Γ is in Γ; this we easily see by differentiating under the integral sign and integrating by parts.

We shall call a function f *radial* if it only depends on $|x|$. Fourier transforms of radial functions are radial (see [1] page 67). By a *corradial* function on the other hand we shall mean a function which is orthogonal to

all radial functions, i. e. such that $f \cdot g = 0$ for all radial g. The fact just quoted clearly implies that Fourier transforms of corradial functions are corradial.

Homogeneous functions satisfying the first condition 1.1 are corradial in an obvious sense, and conversely every corradial homogeneous function satisfies that condition. We shall therefore refer to homogeneous functions satisfying 1.1 as *corradial homogeneous functions.*

The argument which follows is based on a certain representation of homogeneous functions of a given degree.

Suppose that $g(x)$ is a corradial function in Γ. Then $g(0) = 0$ and

$$\int_0^\infty g^\lambda(x)\lambda^{-n-1+r}d\lambda \tag{2.1}$$

converges absolutely for $r > -1$. Moreover it represents a corradial homogeneous function of degree $-r$. Differentiation under the integral sign shows that this function is of C^∞ in $x \neq 0$.

Conversely, every corradial homogeneous function $K(x)$ of degree $-r$ can thus be represented by setting $g(x) = K(x)\rho(|x|)$ where $\rho(t)$ has continuous derivatives of all orders, vanishes in a neighborhood of 0 and ∞ nad such that

$$\int_0^\infty \lambda^{-1}\rho(\lambda)\,d\lambda = 1. \tag{2.2}$$

Let $K(x)$ be corradial homogeneous of degree $-n$ and of C^∞ in $x \neq 0$. Then, if $x \neq 0$, $K_\lambda(x)$ converges to $K(x)$ as $\lambda \to 0$, and it is not difficult to prove (see [2], pp. 89-91) that also $\mathfrak{F}(K_\lambda)$ converges pointwise and boundedly to a limit which we shall denote by $\mathfrak{F}(K)$. Consequently, if $f \varepsilon L^2$, $K_\lambda * f$ converges in mean of order 2, and the Fourier transform of its limit is $\mathfrak{F}(f)\mathfrak{F}(K)$. Thus the Fourier transform $\mathfrak{F}(\mathcal{K})$ of the operator in (1.4) is precisely $\alpha + \mathfrak{F}(K)$.

We shall prove presently that this function is homogeneous of degree zero and of C^∞ in $x \neq 0$, and that conversely every function with such properties is of this form. This will imply immediately that $\mathfrak{A}$ is closed under operator multiplication, as we stated in Theorem 1.

Let $\rho(x)$ be a radial function of C^∞ such that $\rho(0) = 1$, $\rho(x) = 0$ for $|x| \geqq 1$, and let $f(x)$ be any function of C^∞ vanishing outside a bounded set. Then $K_\lambda \cdot \rho = 0$ and

$$\begin{aligned} \mathfrak{F}(f) \cdot \mathfrak{F}(K) &= \lim_{\lambda\to 0} \mathfrak{F}(f) \cdot \mathfrak{F}(K_\lambda) = \lim_{\lambda\to 0}\,(f \cdot K_\lambda) \\ &= \lim_{\lambda\to 0}\,[f - f(0)\rho] \cdot K_\lambda = [f - f(0)\rho] \cdot K, \end{aligned} \tag{2.3}$$

the last integral being absolutely convergent since $f(x)-f(0)\rho(x)$ vanishes at 0. We now represent $K(x)$ by the formula (2.1) with $r=n$ and obtain

$$[f-f(0)\rho]\cdot K=[f-f(0)\rho]\cdot\int_0^\infty\lambda^{-1}g^\lambda d\lambda=\int_0^\infty\lambda^{-1}[f-f(0)\rho]\cdot g^\lambda d\lambda,$$

the change of the order of integration being justified by absolute convergence.

Now since g^λ is corradial and ρ radial we have $g^\lambda\cdot\rho=0$, and since

$$\mathfrak{F}(g^\lambda)=\lambda^n\mathfrak{F}(g)^{\lambda^{-1}},$$

as seen by changing variables in the Fourier integral of g^λ, setting $\lambda^{-1}=\mu$ we may further write

$$(2.4)\quad \begin{aligned}\int_0^\infty\lambda^{-1}[f-f(0)\rho]\cdot g^\lambda d\lambda&=\int_0^\infty\lambda^{-1}(f\cdot g^\lambda)d\lambda=\int_0^\infty\mathfrak{F}(f)\cdot\mathfrak{F}(g^\lambda)\lambda^{-1}d\lambda\\&=\int_0^\infty\mathfrak{F}(f)\cdot\mathfrak{F}(g)^\mu\mu^{-n-1}d\mu=\mathfrak{F}(f)\cdot\int_0^\infty\mathfrak{F}(g)^\mu\mu^{-n-1}d\mu\end{aligned}$$

changes of the order of integration being again justified by the absolute convergence of the integrals involved.

From the equality of the left side of (2.3) and the right side of (2.4) we conclude that if

$$(2.5)\qquad K(x)=\int_0^\infty\lambda^{-1}g^\lambda(x)d\lambda,$$

then

$$(2.6)\qquad \mathfrak{F}(K)=\int_0^\infty\mathfrak{F}(g)^\lambda\lambda^{-n-1}d\lambda.$$

Since these integrals represent the most general corradial homogeneous functions of degrees $-n$ and 0 respectively, of C^∞ in $x\neq0$, we have proved *that* $\mathfrak{F}(\mathcal{K})$ is corradial homogeneous of degree zero and of C^∞ in $x\neq0$, and that conversely every function with these properties is an $\mathfrak{F}(\mathcal{K})$.

We now pass to the proof of (1.9). For this purpose we assume that $K(x)$ is represented as in (2.5), and that $|x|=1$. Then

$$\begin{aligned}|K(x)|&=\left|\int_0^\infty g(\lambda x)\lambda^{n-1}d\lambda\right|\\&\leqq\sup|g(y)|\int_0^1\lambda^{n-1}d\lambda+\sup|g(y)|\,|y|^{n+1}\int_1^\infty\lambda^{-2}d\lambda,\end{aligned}$$

and we only have to estimate $\sup|g(y)|$ and $\sup|g(y)|\,|y|^{n+1}$ in terms of $\mathfrak{F}(K)$.

Let $\hat{g}$ denote the Fourier transform of g, and let $\xi_1,\xi_2,\cdots,\xi_n$ be the coordinates of x, and $\eta_1,\cdots,\eta_n$ those of y. Then

$$g(x)=\int e^{2\pi i(x\cdot y)}\hat{g}(y)\,dy,$$

$$\xi_k^{\,n+1}g(x)=(2\pi i)^{-n-1}\int e^{2\pi i(x\cdot y)}(\partial^{n+1}/\partial\eta_k^{\,n+1})\hat{g}(y)\,dy.$$

By Hölder's inequality $|x|^{n+1}\leqq n^{\frac{1}{2}(n-1)}\sum|\xi_i|^{n+1}$, and thus

$$|g(x)|\,|x|^{n+1}\leqq(2\pi)^{-n-1}n^{\frac{1}{2}(n-1)}\sum_k\int|(\partial^{n+1}/\partial\eta_k^{\,n+1})\hat{g}(y)|\,dy,$$

$|g(x)|\leqq\int|\hat{g}(y)|\,dy$, and it only remains to estimate the integral on the right in terms of $\mathfrak{F}(K)$.

For this purpose we set, as we may, $\hat{g}(y)=\mathfrak{F}(K)\rho(|y|)$, where $\rho(\lambda)$ is of C^∞, vanishes outside $1\leqq\lambda\leqq 2$ and satisfies (2.2). This function we choose once for all independently of the particular kernel K under consideration. This makes it clear that $\hat{g}(y)$ and its derivatives can be estimated in terms of the derivatives of $\mathfrak{F}(K)$ and $\mathfrak{F}(K)$ itself. Furthermore $\hat{g}(y)$ vanishes outside $1\leqq|y|\leqq 2$, and this fact makes it possible to estimate the integral above in terms of $\beta[\mathfrak{F}(K)]$. Collecting estimates we obtain (1.9).

Finally, from (1.9) we readily obtain

(2.8) $$\|\mathcal{K}\|_p\leqq A\beta[\mathfrak{F}(\mathcal{K})]$$

for any operator $\mathcal{K}$ in $\mathfrak{A}$, A being a constant independent of p and $\mathcal{K}$, but not necessarily the same as in (1.9).

3. In this section we prove (1.6) for operators in $\mathfrak{A}$.

Let K and H be two corradial homogeneous functions of degree $-n$ of C^∞ in $x\neq 0$, and consider

(3.1) $$K*H_\lambda=\lim_{\mu\to 0}K_\mu*H_\lambda.$$

As $\lambda\to 0$, this function converges pointwise for $x\neq 0$, and its limit J is a homogeneous function of degree $-n$. On the other hand, the Fourier transform of (3.1) converges to $\mathfrak{F}(K)\mathfrak{F}(H)$. We will show that J is corradial and that

(3.2) $$\mathfrak{F}(K)\mathfrak{F}(H)=\alpha+\mathfrak{F}(J),$$

where α is a constant, and that

(3.3) $$\|\mathcal{J}\|_p\leqq A_p\|\mathcal{K}\|_p\|\mathcal{H}\|_p,$$

(3.4) $$|\alpha|\leqq A_p\|\mathcal{K}\|_p\|\mathcal{H}\|_p,$$

$\mathcal{H}$, $\mathcal{J}$, $\mathcal{K}$ being the convolution operators with kernels H, J, K, respectively.

First we easily see that $H_\lambda = (H_1)^{\lambda^{-1}}$; $K * H_\lambda = (K * H_1)^{\lambda^{-1}}$. Thus $K * H_\lambda - J_\lambda = (K * H_1 - J_1)^{\lambda^{-1}}$. Since H_1 is in L^2 the same holds for $K * H_1$. It follows that $K * H_1 - J_1$ is integrable over bounded sets. On the other hand it is not difficult to see that $K * H_1 - J_1$ is of order $|x|^{-n-1}$ as $|x| \to \infty$. Hence $K * H_1 - J_1$ is absolutely integrable.

Let now f be a function in Γ. A change of variables gives

$$(K * H_\lambda - J_\lambda) \cdot f = (K * H_1 - J_1)^{\lambda^{-1}} \cdot f = (K * H_1 - J_1) \cdot \lambda^{-n} f^\lambda$$

As $\lambda \to 0$, $\lambda^{-n} f^\lambda$ tends to $f(0)$ while remaining bounded, and this implies that $(K * H_\lambda - J_\lambda) \cdot f$ converges as $\lambda \to 0$. Now

$$K * H_\lambda \cdot f = \mathfrak{F}(K)\, \mathfrak{F}(H_\lambda) \cdot \mathfrak{F}(f)$$

also converges as $\lambda \to 0$, and consequently the same holds for $J_\lambda \cdot f$. But if $f(0) \neq 0$, $J_\lambda \cdot f$ cannot converge unless J is corradial. Hence J is corradial and $\mathfrak{F}(J_\lambda)$ converges boundedly to a limit $\mathfrak{F}(J)$.

Suppose now that $f(0) = 0$. Then

$$\begin{aligned}[\mathfrak{F}(K)\mathfrak{F}(H) - \mathfrak{F}(J)] \cdot \mathfrak{F}(f) &= \lim_{\lambda \to 0} [\mathfrak{F}(K)\mathfrak{F}(H_\lambda) - \mathfrak{F}(J_\lambda)] \cdot \mathfrak{F}(f) \\ &= \lim_{\lambda \to 0} (K * H_\lambda - J_\lambda) \cdot f.\end{aligned}$$

Since $f(0) = 0$, the last limit is zero, as we pointed out above. Consequently, if $g = \mathfrak{F}(f)$ we have

$$[\mathfrak{F}(K)\mathfrak{F}(H) - \mathfrak{F}(J)] \cdot g = 0$$

for any $g \in \Gamma$ with vanishing integral, and this is possible only if $\mathfrak{F}(K)\mathfrak{F}(H) - \mathfrak{F}(J)$ is a constant.

Next we estimate $\| \mathcal{J} \|_p$. First we note that, on account of homogeneity,

$$(3.5) \qquad [\int |J_\lambda(x)|^p\, dx]^{1/p} = [n(p-1)\lambda^{(p-1)n}]^{-1/p} \| \mathcal{J} \|_p,$$

and similarly for H and K. Next, for $|x| \geqq 2$ we have

$$\begin{aligned}J_2(x) &= [(K - K_1) + K_1] * [(H - H_1) + H_1] \\ &= (K - K_1) * (H - H_1) + (K - K_1) * H_1 + (H - H_1) * K_1 + H_1 * K_1,\end{aligned}$$

and since the first term in the last sum vanishes for $|x| \geqq 2$, we see that

$$J_2(x) = K * H_1 + H * K_1 - H_1 * K_1$$

for $|x| \geqq 2$. Now (3.5) and (1.3) applied to this inequality yield (3.3). And from (1.8), (3.3) and (3.2) we easily derive (3.4).

It is clear that (3.3) and (3.4) imply (1.6).

4. The extension of (1.6) to operators in $\mathcal{A}_p$ is straightforward.

Given two operators $\mathcal{H}$ and $\mathcal{K}$ in $\mathcal{A}_p$, we take two sequences of operators $\mathcal{H}_n$, $\mathcal{K}_n$ in $\mathcal{A}$ such that

$$\| \mathcal{H}_n - \mathcal{H} \|_p \to 0; \qquad \| \mathcal{K}_n - \mathcal{K} \|_p \to 0.$$

Then from the validity of (1.6) for operators in $\mathcal{A}$ it follows that $\mathcal{H}_n\mathcal{K}_n$ is a Cauchy sequence in $\mathcal{A}_p$, and therefore converges to a limit $\mathcal{J}$ in $\mathcal{A}_p$ for which the inequality $\| \mathcal{J} \|_p \leqq A_p \| \mathcal{H} \|_p \| \mathcal{K} \|_p$ holds. Consequently, if we show that $\mathcal{J} = \mathcal{H}\mathcal{K}$ we will have shown that $\mathcal{A}_p$ is closed under multiplication (composition) and that (1.6) holds for the product.

Consider (1.3) and (1.5). Assuming, as we may, that $A_{r,p} \geqq 1$, we see that $\mathcal{K}(f) = \alpha f + \tilde{f}$ satisfies $\| \mathcal{K}(f) \|_r \leqq A_{r,p} \| \mathcal{K} \|_p \| f \|_r$. Consequently the operator norm of $\mathcal{K}$ as an operator in L^r, which is defined as

$$\sup_f \| \mathcal{K}(f) \|_r / \| f \|_r,$$

is dominated by $A_{r,p} \| \mathcal{K} \|_p$. Since $\mathcal{H}_n \to \mathcal{H}$ and $\mathcal{K}_n \to \mathcal{K}$ in $\mathcal{A}_p$, the same holds in the operator topology, and consequently $\mathcal{H}_n\mathcal{K}_n \to \mathcal{H}\mathcal{K}$ in the operator topology. On the other hand, $\mathcal{H}_n\mathcal{K}_n \to \mathcal{J}$ in $\mathcal{A}_p$, and consequently the same holds in the operator topology. Hence $\mathcal{J} = \mathcal{H}\mathcal{K}$ and the proof is completed.

This also completes the proof of Theorem 2 since the fact that the Fourier transforms of operators in $\mathcal{A}_p$ are continuous homogeneous functions of degree zero follows readily from (1.8) and the fact that $\mathcal{A}$ is dense in $\mathcal{A}_p$ and its elements have continuous Fourier transforms.

5. We now proceed to prove Theorems 3 and 4.

We might observe here that if we knew already that every maximal ideal in $\mathcal{A}_p$ is the set of all operators whose Fourier transforms vanish at a point of the unit sphere Σ ($|x| = 1$), then Theorems 3 and 4 would merely be standard facts from Banach Algebras. In our present setup though we can prove Theorems 3 and 4 directly with comparatively little additional effort and obtain the structure of the maximal ideals as a consequence.

For simplicity of notation we shall denote the Fourier transform of an operator $\mathcal{H}$ in $\mathcal{A}_p$ by h. The symbol $\| h \|_p$ will now stand for $\| \mathcal{H} \|_p$, and $\beta(h)$ will denote, as in Theorem 1 or Section 2, the least upper bound for the absolute value of h and its derivatives of order $n+1$, evaluated in $|x| \geqq 1$. Occasionally, instead of working with homogeneous functions of degree zero we shall work with their restrictions to the unit sphere Σ.

Let $g(x)$ be a function on Σ, and suppose that for each x_0 there is a

neighborhood N_{x_0} of x_0 and an operator in $\mathcal{H}_{x_0}$ whose Fourier transform h_{x_0} (restricted to Σ) coincides with $g(x)$ in N_{x_0}. Let N_{x_i}, $i=1,2,\cdots$ be a finite collection of such neighborhoods covering Σ. Let further $k_i \geqq 0$ be functions in C^∞, each vanishing outside N_{x_i} and such that $\Sigma\, k_i(x) > 0$. Then

$$k'_i(x) = k_i(x)\,[\sum_j k_j(x)]^{-1}$$

is also in C^∞ and vanishes outside N_{x_i}. Furthermore $\Sigma\, k'_i(x) = 1$, and consequently

$$g(x) = \Sigma\, g(x) k'_i(x) = \Sigma\, h_{x_i}(x) k'_i(x),$$

since $h_{x_i}(x) = g(x)$ wherever $k'_i(x) \neq 0$. Since $k'_i(x)$ is a restriction to Σ of a homogeneous function of degree zero of C^∞ in $x \neq 0$, which is in turn the Fourier transform of an operator $\mathcal{K}_i$ in $\mathcal{A}$, the last expression on the right is precisely the restriction to Σ of the Fourier transform of $\Sigma\, \mathcal{H}_{x_i}\mathcal{K}_i$, and Theorem 3 is thus established.

To prove Theorem 4 we begin by observing that, as an easy computation shows, if $\alpha > |h(x)|$ and $h(x)$ is in C^∞, then $\beta(h^k) = O(\alpha^k)$ as $k \to \infty$. Consequently it follows from (2.8) that $\| h^k \|_p = O(\alpha^k)$.

We now extend this result to the Fourier transform $h(x)$ of an arbitrary operator $\mathcal{H}$ in $\mathcal{A}_p$. Given such an $h(x)$ and $\alpha > |h(x)|$, we take α_0 so that $\alpha > \alpha_0 > |h(x)|$, and $h_0(x)$ in C^∞ so that $A_p \| h - h_0 \|_p + \alpha_0 < \alpha$ and $|h_0(x)| < \alpha_0$. Then, by (1.6),

$$\| h^k \|_p = \| [(h-h_0) + h_0]^k \|_p \leqq \sum_{i=0}^{k} \binom{k}{i} \| (h-h_0)^i h_0^{k-i} \|_p$$

$$= O[\sum_{i=0}^{k} A_p{}^i \binom{k}{i} \| h - h_0 \|_p{}^i \alpha_0{}^{k-i}] = O[(A_p \| h - h_0 \|_p + \alpha_0)^k] = O(\alpha^k).$$

Let now $F(z) = \Sigma\, a_k (z - z_0)^k$ be analytic in $|z - z_0| < 2\epsilon$, $h(x)$ the Fourier transform of an operator in $\mathcal{A}_p$, and $h(x_0) = z_0$. If we show that $g(x) = F[h(x)]$ coincides with the Fourier transform of an operator in $\mathcal{A}_p$ in some neighborhood of x_0, a repeated application of this result and Theorem 3 will yield Theorem 4. For this purpose we take $0 \leqq k(x) \leqq 1$ homogeneous of degree zero, equal to 1 in a neighborhood of x_0 and vanishing wherever $|h(x) - h(x_0)| \geqq \epsilon$, and define

$$h'(x) = h(x_0) + k(x)[h(x) - h(x_0)].$$

Clearly we have $|h'(x) - h(x_0)| < \epsilon$ and $h'(x) = h(x)$ in a neighborhood of x_0. The series

$$F[h'(x)] = \Sigma\, a_k [h'(x) - h(x_0)]^k$$

coincides with $F[h(x)]$ in a neighborhood of x_0. But since

$$\|[h'(x)-h(x_0)]^k\|_p = O(\epsilon^k),$$

the corresponding series of operators converges in $\mathfrak{A}_p$, and the Fourier transform of its sum is precisely $F[h'(x)]$. Theorem 4 is thus established.

Regarding the Corollary to Theorem 4 we observe that if $h(x)$ does not vanish then $h^{-1}(x)$ satisfies the conditions of Theorem 4 and consequently it is the Fourier transform of an operator in $\mathfrak{A}_p$.

To determine the structure of the maximal ideals in $\mathfrak{A}_p$ we observe that if $h_i = \mathfrak{F}(\mathfrak{H}_i)$ and the $\mathfrak{H}_i$ belong to a proper ideal I in $\mathfrak{A}_p$, the $h_i(x)$ must be necessity have a common zero. For otherwise there would exist a finite number of such $h_i(x)$ without common zero, and the function $h(x) = \Sigma h_i \bar{h}_i > 0$ would be the Fourier transform of an invertible operator in I, and I would not be a proper ideal. Consequently a maximal ideal in $\mathfrak{A}_p$ consists of all operators whose Fourier transform vanish at a point of Σ, and conversely.

MASSACHUSETTS INSTITUTE OF TECHNOLOGY AND
THE UNIVERSITY OF CHICAGO.

REFERENCES.

[1] S. Bochner and K. Chandrasekharan, "Fourier transforms," *Annals of Mathematics Studies*, 19 (1949).

[2] A. P. Calderón and A. Zygmund, "On the existence of certain singular integrals," *Acta Mathematica*, vol. 88 (1952), 85-139.

[3] ———, "On singular integrals," *American Journal of Mathematics*, vol. 78 (1956), pp. 289-309.

ON THE VALUATIONS CENTERED IN A LOCAL DOMAIN.*[1]

By Shreeram Abhyankar.

It is well known that if v is a real zero dimensional valuation of an algebraic function field K/k of n variables and if r is the rational rank of v then the following is true: $r \leqq n$ and if $r = n$ then the value group of v a direct sum of n cyclic groups; see Theorem 1 of [8]. In Section 1, we prove a generalization of this theorem to arbitrary valuations (real or not) centered in an abstract local domain. Namely, let R be a local domain of dimension n with maximal ideal M and quotient field K. Let v be an arbitrary valuation of K having center M in R. Let ρ be the rank and r the rational rank of v, and let d be the transcendence degree of the residue field D of v over R/M. Then we have the following: (1) $d + r \leqq n$, (2) if $d + r = n$ then the value group of v is a direct sum of r cyclic groups and $D/(R/M)$ is finitely generated, (3) if $d + \rho = n$ then v is discrete and $D/(R/M)$ is finitely generated.

In Section 2, we generalize some theorems concerning quadratic sequences of quotient rings on a nonsingular algebraic surface to abstract regular local domains. In Theorem 2, we prove that if f is a nonzero element in a two dimensional regular local domain (R, M) with quotient field K and if v is a valuation of K having center M in R then there exists a quadratic transform (R^*, M^*) of R along v and a basis (x^*, y^*) of M^* such that $f = x^{*a}y^{*b}d$ where a and b are nonnegative integers and d is a unit in R^*, (if v is nonreal, then we assume that R is either algebraic or absolute; for definitions see the beginning of Section 1). The special case of Theorem 2 when R is the quotient ring of a point on an algebraic surface and when v is real plays an important role in the proof of the local uniformization theorem (see Proposition 3 of [2] and Lemma 11.2 of [11]). In Theorem 3 we generalize Zariski's factorization theorem on birational transformations between algebraic surfaces to abstract regular two dimensional local domains. As an incidental remark, we show in Proposition 3 that if (R, M) is a regular n dimensional local domain and if w is a valuation of the quotient field of R with center M in R and of R-dimension $n-1$ then R_w/M_w is a purely transcendental exten-

* Received August 1, 1955.

[1] This work was supported by a research project at Harvard University, sponsored by the National Science Foundation.

sion of a finitely generated extension of R/M. It is a consequence of this proposition that if w is a prime divisor of the second kind having center at a simple subvariety of an algebraic variety then the residue field of w is the function field of a ruled variety. We hope to use the results of this paper in the problem of local uniformization on *absolute* surfaces which we are planning to study.

1. Classification of valuations. We start with some remarks about ordered abelian groups. Let G be an ordered abelian group. Let us recall that the rational rank of G is defined to be the maximum cardinal number (finite or infinite) of a subset H of G such that any finite number $h_1, h_2, \cdots, h_n$ of distinct elements of H are rationally independent, i.e., if $m_1h_1 + m_2h_2 + \cdots + m_nh_n = 0$ where the m_i are integers then $m_1 = m_2 = \cdots = m_n = 0$. We observe that if G has finite rational rank r then we can find r elements $g_1, g_2, \cdots, g_r$ in G such that for any element g in G there exist integers $m_1, m_2, \cdots, m_r$ and a nonzero integer m for which $mg = m_1g_1 + m_2g_2 + \cdots + m_rg_r$ (we call $\{g_1, g_2, \cdots, g_r\}$ a rational basis of G); and conversely if there exist r elements $g_1, g_2, \cdots, g_r$ in G with this property then G is of finite rational rank $n \leqq r$. We assert that the following two conditions are equivalent: (1) G is of finite rational rank r. (2) G is of finite rank ρ and if $(0) = G_0 < G_1 < G_2 < \cdots < G_\rho = G$ is the sequence of isolated subgroups of G then G_i/G_{i-1} is of finite rational rank r_i for $i = 1, 2, \cdots, \rho$. Furthermore, when one and hence both of the above conditions are satisfied, we have the equation: $r = r_1 + r_2 + \cdots + r_\rho$. This is easily verified by writing the elements of G as lexicographically ordered ρ-tuples of real numbers[2] (roughly speaking, if we put together rational bases of G_1/G_0, $G_2/G_1, \cdots, G_\rho/G_{\rho-1}$ then we get a rational basis of G). Given an additive group F of real numbers, we shall say that F is an *integral direct sum*, if F is of finite rational rank a and if we can find elements $f_1, f_2, \cdots, f_a$ in F such that any element f in F can be expressed as: $f = m_1f_1 + m_2f_2 + \cdots + m_af_a$ where the m_i are integers. Again, given an ordered abelian group G, we shall say that G is an integral direct sum if G is of finite rational rank r and if G_i/G_{i-1} is an integral direct sum for $i = 1, 2, \cdots, \rho$ where $(0) = G_0 < G_1 < G_2 < \cdots < G_\rho = G$ is the sequence of isolated subgroups of G. One can easily show that G is an integral direct sum if and only if G is of finite rational rank r and if G contains elements $g_1, g_2, \cdots, g_r$ such that any element g of G can be expressed as: $g = m_1g_1 + m_2g_2 + \cdots + m_rg_r$ where the m_i

[2] See the Appendix.

are integers ($g_1, g_2, \cdots, g_r$ are then necessarily rationally independent; we call $\{g_1, g_2, \cdots, g_r\}$ an integral basis of G). We observe that *G is discrete if and only if G is an integral direct sum and the rational rank of G equals the rank of G.* Finally, note that if H is a subgroup of G of finite index then the rank (respectively, the rational rank) of H equals the rank (respectively, the rational rank) of G and H is an integral direct sum (respectively, discrete) if and only if G is an integral direct sum (respectively, discrete).

We shall consistently use, in this paper, the following *notations.* Let v be a valuation of a field K. We shall denote by R_v the valuation ring of v and by M_v the maximal ideal in R_v. By the rank (respectively, rational rank) of v we shall mean the rank (respectively, rational rank) of the value group of v, also we shall say that v is an integral direct sum if the value group of v is an integral direct sum, and so on. Unless otherwise stated, we shall exclude trivial valuations. When we do want to talk of a trivial valuation, we shall assign to it the zero rank and the zero rational rank, since its value group consists of the zero element alone, and for a trivial valuation we shall consider the designations "integral direct sum" and "discrete" as trivially valid. If v^* is a valuation of a field K^* and v its restriction to a subfield K, then by the v-dimension of v^* is meant the transcendence degree of R_{v^*}/M_{v^*} over R_v/M_v; if v is trivial (over K) then we have that the v-dimension of v^* equals the K-dimension of v, i.e., the transcendence degree of R_{v^*}/M_{v^*} over K. If R is a local ring and M its maximal ideal, we shall indicate this by saying that (R, M) is a local ring. If a local ring R is the quotient ring of an irreducible subvariety of an algebraic variety, then we shall say that R is *algebraic.* By Q and Z we shall denote the field of rational numbers and the domain of ordinary integers respectively. If a local ring R is the quotient ring of a domain A finitely generated over Z (i.e., $A = Z[x_1, x_2, \cdots, x_n]$ with x_i in A) with respect to a prime ideal in A, then we shall say that R is *absolute.* If (R, M) is a local domain and v a valuation of the quotient field of R having center in R (i.e., such that $R_v \supset R$ and $R \cap M_v = M$), then by the R-dimension of v is meant the transcendence degree of R_v/M_v over R/M. Recall that if R is a (commutative) ring and P a nonzero prime ideal in R, then by the dimension of P is meant the maximum value of n such that there exists a strictly ascending chain $P = P_0 < P_1 < P_2 < \cdots < P_n < R$ of prime ideals P_i, and by the rank of P is meant the maximum value of n such that there exists a strictly descending chain $P = P_0 > P_1 > P_2 > \cdots > P_n \supset (0)$ of prime ideals P_i (where P_n is allowed to be the zero ideal if this be prime). Finally, a given integral domain F will be called normal if F is integrally closed in its quotient field.

LEMMA 1. *Let K be a field and K^* an extension of K of finite transcendence degree s. Let v^* be a valuation of K^* and let v be the K-restriction of v^* where we allow v to be trivial. Let d be the v-dimension of v^*. Let r and r^* be the rational ranks of v and v^* and let ρ and ρ^* be the ranks of v and v^* respectively. Then we have the following*: (1) *If r^* is finite, then $r^* + d \leqq r + s$.* (2) *If v is an integral direct sum, if K^*/K is finitely generated, and if $r^* + d = r + s$, then v^* is an integral direct sum and R_{v^*}/M_{v^*} is finitely generated over R_v/M_v (note that $R_v/M_v = K$ if v is trivial).* (1*) *If ρ is finite, then $\rho^* + d \leqq \rho + s$.* (3) *If v is discrete, if K^*/K is finitely generated, and if $\rho^* + d = \rho + s$, then v^* is discrete and R_{v^*}/M_{v^*} is finitely generated over R_v/M_v.*

Proof. First assume that r is finite. We begin by proving the weaker inequality:

(A) $$r^* \leqq r + s.$$

Suppose $s = 0$. Given $0 \neq u \,\varepsilon\, K^*$, let $f(X) = A_0X^n + a_1X^{n-1} + \cdots + a_n$, $a_0 = 1$, be the minimal monic polynomial of u over K. Since $f(u) = 0$, there exist distinct integers i and j such that $v^*(a_iu^{n-i}) = v^*(a_ju^{n-j}) \neq \infty$ and hence $v^*(u) = v(a_i/a_j)/(i - j)$, i.e., the value of u depends rationally on the value of $(a_i/a_j) \,\varepsilon\, K$. Therefore $r^* = r = r + s$. Now suppose $s > 0$ and assume that (A) is true for $s - 1$. Let $z_1, z_2, \cdots, z_{s-1}$ be part of a transcendence basis of K^*/K. Let $K_1 = K(z_1, z_2, \cdots, z_{s-1})$, let v_1 be the restriction of v^* to K_1 (v_1 may be trivial), and let r_1 be the rational rank of v_1. By our induction hypothesis, $r_1 \leqq r + s - 1$. If the value of every nonzero element of K^* is rationally dependent on the values of elements of K_1, then $r^* = r_1 \leqq r + s - 1 \leqq r - s$, and we are through. Now suppose that there is a nonzero element z in K^* such that $h = v^*(z)$ does not depend rationally on the values of elements of K_1. Then, by the $s = 0$ case, z is transcendental over K_1. Let $f(X) = f_0 + f_1X + \cdots + f_nX^n$ and $g(X) = g_0 + g_1X + \cdots + g_{n^*}X^{n^*}$ be nonzero elements of $K_1[X]$. Let $a_i = v^*(f_i)$ if $f_i \neq 0$ and $b_i = v^*(g_i)$ if $g_i \neq 0$. Since h depends rationally neither on the a_i nor on the b_i, there exist integers p and q such that $\infty \neq v^*(f_pz^p) < v^*(f_iz^i)$ whenever $i \neq p$ and $f_i \neq 0$, and $\infty \neq v^*(g_qz^q) < v^*(g_iz^i)$ whenever $i \neq q$ and $g_i \neq 0$; i.e., $v^*(f(z)/g(z)) = v^*(f_p/g_q) + (p - q)h$. Thus, *the value of any nonzero element of $K_1(z)$ is of the form $a + mh$ where a is in the value group of v_1 and m is an integer*, i.e., if r_2 is the rational rank of the restriction of v^* to $K_1(z)$ (this restriction may be trivial), then

$$r_2 = r_1 + 1 \leqq r + (s - 1) + 1 = r + s.$$

Since $K^*/K_1(z)$ is an algebraic extension, by the case $s=0$, we have $r^*=r_2\leqq r+s$. Thus the induction is complete and (A) has been proved. *Also observe that if v_2 is the restriction of v^* to $K_2=K_1(z)$, then the residue fields of v_1 and v_2 coincide.* For, in the above notation, since $v_2(f_iz^i)>v_2(f_pz^p)$ whenever $i\neq p$ and $f_i\neq 0$, we must have that $f(z)/(f_pz^p)$ belongs to R_{v_2} and that

$$f(z)/(f_pz^p)=1+\sum_{i\neq p}(f_i/f_p)z^{i-p}\equiv 1 \pmod{M_{v_2}}.$$

Similarly, $g(z)/(g_qz^q)$ belongs to R_{v_2} and $g(z)/(g_qz^q)\equiv 1 \pmod{M_{v_2}}$. Now assume that $f(z)/g(z)$ belongs to R_{v_2}. We want to show that we can find e in R_{v_1} with $f(z)/g(z)\equiv e \pmod{M_{v_2}}$. If $f(z)/g(z)$ belongs to M_{v_2}, we can take $e=0$. Now suppose that $f(z)/g(z)$ does not belong to M_{v_2}, i.e., that $v_2(f(z)/g(z))=0$. Since $v_2(f(z)/g(z))=v_2(f_p/_q)+(p-q)h$ and since h does not depend rationally on $v_2(f_p/g_q)$, we must have $p-q=0$, i.e., that $p=q$ and $v_2(f_p/g_p)=0$. Let $e=f_p/g_p$. Then

$$f(z)/g(z)=(f_p/g_p)(f(z)/f_pz^p)(g_pz^p/g(z))\equiv e \pmod{M_{v_2}}$$

since $f(z)/f_pz^p$ and $g_pz^p/g(z)$ are both congruent to one modulo M_{v_2}. This proves our second italicized assertion.

To prove (1) let us retain our assumption that r is finite, and let $\bar{y}_1,\bar{y}_2,\cdots,\bar{y}_d$ be a transcendence basis of R_{v^*}/M_{v^*} over R_v/M_v and fix y_i in R_{v^*} belonging to the residue class $\bar{y}_i$. Let $K'=K(y_1,y_2,\cdots,y_d)$ and let v' be the restriction of v^* to K'. Given $0\neq f(X_1,X_2,\cdots,X_d)$ in $K[X_1,X_2,\cdots,X_d]$ choose a coefficient q of f having minimum v-value and let $F(X_1,X_2,\cdots,X_d)$ $=(1/q)\ f(X_1,X_2,\cdots,X_d)$. Then all the coefficients of $F(X_1,X_2,\cdots,X_d)$ belong to R_v and at least one of them is equal to 1. Let $\bar{F}(X_1,X_2,\cdots,X_d)$ be the polynomial gotten by reducing the coefficients of $F(X_1,X_2,\cdots,X_d)$ modulo M_v. Since $\bar{F}(X_1,X_2,\cdots,X_d)$ has a coefficient equal to 1 and since $\bar{y}_1,\bar{y}_2,\cdots,\bar{y}_d$ are algebraically independent over R_v/M_v, we must have $\bar{F}(\bar{y}_1,\bar{y}_2,\cdots,\bar{y}_d)\neq 0$, i.e., $v^*(F(y_1,y_2,\cdots,y_d))=0$, i.e., $v^*(f(y_1,y_2,\cdots,y_d))$ $=v(q)\neq\infty$, and hence $f(y_1,y_2,\cdots,y_d)\neq 0$. Thus $y_1,y_2,\cdots,y_d$ are algebraically independent over K and *the value groups of v and v' are identical.* Since the transcendence degree of K^*/K' is $s-d$, (1) follows by applying (A) to K^*/K'. Now let $g(X_1,X_2,\cdots,X_d)$ and $h(X_1,X_2,\cdots,X_d)$ be arbitrary nonzero elements of $K[X_1,X_2,\cdots,X_d]$ and let

$$y=f(y_1,y_2,\cdots,y_d)/g(y_1,y_2,\cdots,y_d).$$

Fix coefficients a and b of g and h respectively having minimum v'-values

and let $p = a/b$. Let $G(X_1, X_2, \cdots, X_d) = (1/a)g(X_1, X_2, \cdots, X_d)$ and $H(X_1, X_2, \cdots, X_d) = (1/b)h(X_1, X_2, \cdots, X_d)$. Then, as above,

$$v'(g(y_1, y_2, \cdots, y_d)/h(y_1, y_2, \cdots, y_d)) = v'(a/b).$$

Hence $y = g(y_1, y_2, \cdots, y_d)/h(y_1, y_2, \cdots, y_d)$ belongs to $R_{v'}$ if and only if $p = a/b$ belongs to $R_{v'} \cap K = R_v$. Now assume that y does belong to $R_{v'}$. Let $\bar{y}$ and $\bar{p}$ be the residue classes modulo $M_{v'}$ containing y and p respectively. Let $\bar{G}(X_1, X_2, \cdots, X_d)$ and $\bar{H}(X_1, X_2, \cdots, X_d)$ be the polynomials obtained respectively from $G(X_1, X_2, \cdots, X_d)$ and $H(X_1, X_2, \cdots, X_d)$ by reducing their coefficients modulo $M_{v'}$. Since $\bar{H}$ has a coefficient equal to one and since $\bar{y}_1, \bar{y}_2, \cdots, \bar{y}_d$ are algebraically independent over R_v/M_v, we have that

$$\bar{y} = \bar{p}\bar{G}(\bar{y}_1, \bar{y}_2, \cdots, \bar{y}_d)/\bar{H}(\bar{y}_1, \bar{y}_2, \cdots, \bar{y}_d).$$

Therefore $R_{v'}/M_{v'} = (R_v/M_v)(\bar{y}_1, \bar{y}_2, \cdots, \bar{y}_d)$, *and hence in particular* $R_{v'}/M_{v'}$ *is finitely generated over* R_v/M_v.

Now assume that v is an integral direct sum, that K^*/K is finitely generated, and that $r^* + d = r + s$. Let K' and v' be as above. Then v and v' have the same value groups, K^*/K' is a finitely generated extension of transcendence degree $e = s - d = r^* - r$, and $R_{v'}/M_{v'}$ is finitely generated over R_v/M_v. Fix an integral basis $t_1, t_2, \cdots, t_r$ of the value group of v'. Let $x_1, x_2, \cdots, x_e$ be a transcendence basis of K^*/K'. Let $K_i' = K'(x_1, x_2, \cdots, x_i)$, $v_i' =$ the restriction of v^* to K_i', and $r_i' =$ the rational rank of v_i'. Since $r^* = r + e$, we must have, in view of (1), $r_i = r_{i-1} + 1$ for $i = 1, 2, \cdots, e$. Let $v^*(x_i) = t_{r+i}$. By applying the first of the above italicized remarks successively to the extensions $K_1'/K', K_1', \cdots, K_e'/K_{e-1}'$, we conclude that for any nonzero element x of K_e' we have

$$v^*(x) = a + m_{r+1}t_{r+1} + m_{r+2}t_{r+2} + \cdots + m_{r^*}t_{r^*}$$

where a is the value of an element of K' and where $m_{r+1}, m_{r+2}, \cdots, m_{r^*}$ are integers; since $a = m_1t_1 + m_2t_2 + \cdots + m_rt_r$ where $m_1, m_2, \cdots, m_r$ are integers, we finally have: $v^*(x) = m_1t_1 + m_2t_2 + \cdots + m_{r^*}t_{r^*}$. Therefore v_e' is an integral direct sum. Since K^*/K_e' is a finite algebraic extension, the value group of v_e' is a subgroup of the value group of v^* of finite index and hence v^* is an integral direct sum. Now by the second italicized remark above, the residue field of v_e' coincides with the residue field of v'. Since the residue field of v' is finitely generated over the residue field of v and since K^*/K_e' is a finite algebraic extension, we conclude that R_{v^*}/M_{v^*} is finitely generated over R_v/M_v. This proves (2).

The proof of (1*) is entirely similar to that of (1). Finally, assume

that v is discrete, $\rho^* + d = \rho + s$, and that K^*/K is finitely generated. The discreteness of v implies that $\rho = r$. Since by (1), $r^* + d \leqq r + s$ and since $r^* \geqq \rho^*$, it follows that $r^* = \rho^*$ and that $r^* + d = r + s$. Hence by (2), v^* is an integral direct sum and R_{v^*}/M_{v^*} is finitely generated over R_v/M_v. Since $r^* = \rho^*$, v^* is discrete. This proves (3).

COROLLARY 1. *In the notation of the above lemma, assume that v is trivial, i.e., that v^* is a valuation of K^*/K. Then:* (1) $\rho^* + d \leqq r^* + d \leqq s$. *Furthermore, if K^*/K is finitely generated (i.e., if K^*/K is an algebraic function field of dimension s), then we have the following:* (2) *If $r^* + d = s$ then v^* is an integral direct sum and R_{v^*}/M_{v^*} is finitely generated over K.* (3) *if $\rho^* + d = s$ then v^* is discrete and R_{r^*}/M_{r^*} is finitely generated over K.* (4) *If $d = s - 1$ then v^* is real discrete and R_{v^*}/M_{v^*} is finitely generated over K.*

PROPOSITION 1. *Let K be a field which is of finite transcendence degree over its prime field P. Let n be the absolute dimension*[3] *of K and let v be a valuation of K. Let r be the rational rank of v, let ρ be the rank of v, and let d be the absolute dimension of R_v/M_v. Then:* (1) $r + d \leqq n$. *Furthermore, if K is finitely generated over P then we have the following:* (2) *If $r + d = n$ then v is an integral direct sum and R_v/M_v is finitely generated over its prime field.* (3) *If $\rho + d = n$ then v is discrete and R_v/M_v is finitely generated over its prime field.* (4) *If $d = n - 1$ then v is real discrete and R_v/M_v is finitely generated over its prime field.*

Proof. Let w be the restriction of v to P where w may be trivial. If P is of nonzero characteristic then w is trivial, $n =$ the transcendence degree of K over P, and $R_w/M_w = P$; and hence the proposition follows from the previous corollary. Now assume that P is of characteristic zero, i.e., that $P = Q$ (the field of rational numbers). Let D be the w-dimension of v, d^* the absolute dimension of R_w/M_w, and r^* the rational rank of w. If w is trivial, then $r^* = 0$ and $R_w/M_w = Q$, i.e., $d^* = 1$ and hence $r^* + d^* = 1$. If w is nontrivial, then w must be the real discrete valuation given by a prime number p of Z and hence $r^* = 1$ and $R_w/M_w = Z/(pZ)$, i.e., $d^* = 0$ and hence $r^* + d^* = 1$. Thus in both the cases, w is discrete, $r^* + d^* = 1$, and R_w/M_w is its own prime field. Hence $r + D \leqq r^* + (n - 1)$ if and only if $r + (d - d^*) \leqq r^* + (n - 1)$, i.e., if and only if

$$r + d \leqq (d^* + r^*) + (n - 1) = 1 + (n - 1) = n.$$

[3] Let D be an integral domain with quotient field K such that K is of finite transcendence degree s over its prime subfield. Then we define the absolute dimension of D to be s or $s + 1$ according as K is of nonzero or zero characteristic respectively.

Thus $r + D \leqq r^* + (n-1)$ if and only if $r + d \leqq n$ and $r + D = r^* + (n-1)$ if and only if $r + d = n$. Hence (1), (2), (3) follow respectively from parts (1), (2), (3) of Lemma 1 applied to the extension K/Q. Finally, (4) follows from (1) and (3).

DEFINITION 1. *In the notation of Proposition* 1, *if* $d = n - 1$ *then we shall say that* v *is a prime divisor of* K.

Remark 1. Let K be a field of absolute dimension two which is finitely generated over its prime field P of characteristic p; we can express this by saying that *K is the function field of an absolute surface of characteristic p.* Let v be a valuation of K. It follows by Proposition 1 that v is of one of the following four types: (1) v is a prime divisor of K, i. e., v is real discrete and R_v/M_v is of absolute dimension one. If $p \neq 0$ then R_v/M_v is an algebraic function field of one variable over P. If $p = 0$ then R_v/M_v is either a finite algebraic extension of Q or it is an algebraic function field of dimension one over a prime field of nonzero characteristic. In either case, R_v/M_v is the function field of an *absolute curve.* (2) v is discrete of rank two. (3) v is real of rational rank one, but v is not a divisor of K. (4) v is real of rational rank two; in this case v is necessarily an integral direct sum. In cases (2), (3), and (4), R_v/M_v is an algebraic extension of a prime field of characteristic $q \neq 0$ and hence R_v/M_v is perfect; if $p \neq 0$ then $q = p$; in cases (2) and (4), R_v/M_v, being finitely generated over its prime field, is necessarily finite.

If $p \neq 0$ then K/P is an algebraic function field of two variables, i. e., an absolute surface (respectively, an absolute n-dimensional variety) of characteristic $p \neq 0$ is simply an algebraic surface (respectively, an algebraic n-dimensional variety) over the prime field of characteristic p; in this case the above classification of valuations is well known. The present remark signifies that a parallel situation prevails for absolute surfaces of characteristic zero, i. e., algebraic curves over Q considered as surfaces.

PROPOSITION 2. *Let (R, M) be a local domain of dimension n with quotient field K. Let v be a real valuation of K with center M in R. Let d be the R-dimension of v and let r be the rational rank of v. Then $d + r \leqq n$. Furthermore, if $d + r = n$ then v is an integral direct sum and R_v/M_v is finitely generated over R/M.*

We shall divide the proof of this proposition into several lemmas.

LEMMA 2. *Proposition 2 is true in case R is the power series ring $k[[x_1, x_2, \cdots, x_n]]$ in n variables $x_1, x_2, \cdots, x_n$ over a field k.*

Proof. Let

$$K^* = k(x_1, x_2, \cdots, x_n), \quad A = k[x_1, x_2, \cdots, x_n], \quad P = (x_1, x_2, \cdots, x_n)A,$$

$R^* = A_P$, $M^* = PR^*$, and $v^* =$ the restriction of v to K^*. Then (R, M) is the completion of (R^*, M^*) and hence by Lemma 12 of [2], v^* has center M^* in R^* and v^* has the same residue field and value group as v. Since $R/M = R^*/M^*$, d is also the r^*-dimension of v^*. Since $R^*/M^* = k$, the present lemma follows by applying Lemma 1 to the extension K^*/k.

Lemma 3. *Proposition 2 is true in case R is the power series ring $R_w[[x_1, x_2, \cdots, x_{n-1}]]$ in $n-1$ variables $x_1, x_2, \cdots, x_{n-1}$ over a complete valuation ring R_w of a real discrete valuation w such that R_w is of characteristic zero and $M_w = pR_w$ where p is an ordinary prime number.*

Proof. Let k be the quotient field of R_w, $K^* = k(x_1, x_2, \cdots, x_{n-1})$, $A = R_w[x_1, x_2, \cdots, x_{n-1}]$, $P = A \cap M$, $R^* = A_P$, $M^* = PR^*$, and $v^* =$ the K^*-restriction of v. Then it is easily verified that $M^* = (x_1, x_2, \cdots, x_{n-1})R^*$. that (R^*, M^*) is a regular local domain of dimension n, and that (R, M) is the completion of (R^*, M^*); [Proof: $M = (x_1, x_2, \cdots, x_{n-1}, p)R$, $M^i \cap R^* = M^{*i}$ and hence R and R^* are concordant, and so on]. By the argument used in the proof of Lemma 12 of [2], it follows that v^* has center M^* in R^* and v^* has the same value group and residue field as v. Since $R/M = R^*/M^*$, d must also be the R^*-dimension of v^*. Since $M_{v^*} \cap R_w = M^* \cap R_w = M_w$ and $R^*/M^* = R_w/M_w$ (Lemma 5 of [3]), since $R_w \subset R_{v^*}$ and since R_w is a maximal subring of k, it follows that w is the k-restriction of v^* and that d is also the w-dimension of v^*. The present lemma now follows by applying Lemma 1 to the extension K^*/k.

Lemma 4. *Proposition 2 is true if R is an unramified complete regular local ring.*

Proof. By Theorem 15 of [3], R is isomorphic either to the ring described in Lemma 2 or to the ring described in Lemma 3. Therefore, Lemma 4 follows from Lemmas 2 and 3.

Lemma 5. *Proposition 2 is true if R is complete.*

Proof. By Theorem 16 of [3], R is a finite module over a subring S which is an unramified complete regular local ring having the same dimension and residue field as R. Hence Lemma 5 follows from Lemma 4 by observing that valuations do not change their character in passing to a finite algebraic extension.

LEMMA 6. *Proposition 2 is true.*

Proof. Since v is real, we can employ the technique used in the proof of Theorem 1 of [13] to pass from the case of complete local domains to the general case; Lemma 6 then follows from Lemma 5.

THEOREM 1. *Let (R, M) be a local domain of dimension n with quotient field K and let v be a valuation of K with center M in R. Let d, r, and ρ be respectively the R-dimension, the rational rank, and the rank of v. Then: (1) $d + r \leqq n$. (2) If $d + r = n$ then v is an integral direct sum and R_v/M_v is finitely generated over R/M. (3) If $d + \rho = n$ then v is discrete and R_v/M_v is finitely generated over R/M. (4) If $d = n - 1$ then v is real discrete and R_v/M_v is finitely generated over R/M.*

Proof. We observe that (3) and (4) follow at once from (1) and (2), and we proceed to prove (1) and (2) by applying induction to n. Suppose first that $n = 1$. Let D be the integral closure of R in K. Then D is a Dedekind domain (§ 39 of [5]) and hence we must have $R_v \supset D$ and $M_v \cap D$ $=$ a minimal prime idea H in D, and hence $D_H \subset R_v$. Since D_H is a real discrete valuation ring, it is a maximal subring of K and hence $D_H = R_v$, i.e., v is real discrete. Hence, for $n = 1$, (1) and (2) follow from Proposition 2. Now let $n > 1$ and assume that (1) and (2) are true for all smaller values of n. If v is real then (1) and (2) follow from Proposition 2. So asume that v is nonreal. Given a nonzero prime ideal E in R_v, $(R_v)_E$ is the valuation ring of a valuation of K (see [4]), and hence $E \cap R$ must be a nonzero prime ideal in R (since K is the quotient field of R). Since the set T of nonzero prime ideals in R is simply ordered by inclusion [4], the set T^* of the R-contractions of the members of T is also simply ordered by inclusion. Since any chain of prime ideal in R is of length at most n, T^* must be a finite set. Therefore $P = \bigcap_{F \varepsilon T^*} F$ is *the* smallest ideal in T^*, and hence P is a nonzero prime ideal. Let $B = \bigcap_{E \varepsilon T} E$. Then $B \cap R = P \neq (0)$ and hence $B \neq (0)$. Therefore $(R_v)_B$ is the valuation ring of a nontrivial valuation w of K. Since B is the minimal nonzero prime ideal of R_v, w must be real. Let $K^* = R_w/M_w$, $\bar{R} = R/P$, $\bar{M} = M/P$, $\bar{K} =$ the quotient field of $\bar{R}$ in K^*, $v^* =$ the valuation of K^* induced by v (v^* is nontrivial), $\bar{v} =$ the restriction of v^* to $\bar{K}$ where $\bar{v}$ may be trivial. Let $S = R_P$ and $N = PS$. Then w has center N in S. Let d' be the S-dimension of w. Then $d' =$ the transcendence degree of $K^*/\bar{K}$, since $\bar{K} = S/N$. Let $\bar{r}$, r^*, and r' be the rational ranks of $\bar{v}$, v^*, and w respectively. We observe that $r = r^* + r'$

and we divide the remaining argument into two cases according as $P=M$ or $P\neq M$.

Case 1, $P=M$. Then $\bar{R}=\bar{K}$, $\bar{M}=(0)$ and $\bar{v}$ is trivial. Also, the $\bar{K}$-dimension of $v^*=$ the R-dimension of $v=d$. Applying Proposition 2 to the real valuation w we get $d'+r'\leqq n$. Applying Corollary 1 of Lemma 1 to the extension $K^*/\bar{K}$ we get $r^*+d\leqq d'$. Therefore $r+d=r^*+r'+d\leqq n$. Now assume that $r+d=n$. Then $d'+r'=n$ and $r^*+d=d'$. Since $d'+r'=n$, it follows by Proposition 2 that w is an integral direct sum and $K^*/\bar{K}$ is finitely generated. Since $r^*+d=d'$ and since $K^*/\bar{K}$ is finitely generated, it follows by Corollary 1 of Lemma 1 that v^* is an integral direct sum and $R_{v^*}/M_{v^*}=R_v/M_v$ is finitely generated over $\bar{K}=R/M$. Since v is composed of v^* and w and since v^* and w are integral direct sums, it follows that v is also an integral direct sum.

Case 2, $P\neq M$. Let $\bar{n}$ and m be the dimensions of the local rings $\bar{R}$ and S respectively. Then $\bar{n}=\dim P$ and $m=\operatorname{rank} P$. Therefore

(A) $$\bar{n}+m\leqq n.$$

Since $m>0$, we must have $0<\bar{n}<n$. Now $\bar{v}$ is a nontrivial valuation of $\bar{K}$ having center $\bar{M}$ in $\bar{R}$. Let $\bar{r}$ be the rational rank of $\bar{v}$ and let $\bar{d}$ be the $\bar{R}$-dimension of $\bar{v}$. Let d^* be the $\bar{v}$-dimension of v^*. Then

(B) $$\bar{d}+d^*=d.$$

Since w is a real valuation with center N in S, it follows by Proposition 2 that

(C) $$r'+d'\leqq m.$$

Applying our induction hypothesis to the valuation $\bar{v}$ having center $\bar{M}$ in $\bar{R}$ we get

(D) $$\bar{r}+\bar{d}\leqq\bar{n}.$$

By Lemma 1,

(E) $$r^*+d^*\leqq\bar{r}+d'.$$

Adding the inequalities (C), (D), and (E) we get

$$r'+\bar{r}+r^*+d'+\bar{d}+d^*\leqq m+\bar{n}+\bar{r}+d',$$

i.e.,

$$r'+r^*+\bar{d}+d^*\leqq m+\bar{n}.$$

Since $r'+r^*=r$ and since by (B), $\bar{d}+d^*=d$, we obtain:

(F) $$r+d\leqq m+\bar{n}.$$

Therefore, in view of (A), we conclude that $r+d \leqq n$. Now assume that $r+d=n$. Then by (A) and (F) we must have $r+d=m+\bar{n}$. Since $r=r'+r^*$ and since by (B), $d=\bar{d}+d^*$, we obtain:

$$r'+\bar{r}+r^*+d'+\bar{d}+d^*=m+\bar{n}+\bar{r}+d'.$$

Therefore by (C), (D), and (E), it follows that:

(C′) $$r'+d'=m.$$

(D′) $$\bar{r}+\bar{d}=\bar{n}.$$

(E′) $$r^*+d^*=\bar{r}+d'.$$

By our induction hypothesis, (D′) implies that $\bar{v}$ is an integral direct sum and that $R_{\bar{v}}/M_{\bar{v}}$ is finitely generated over $\bar{R}/\bar{M}=R/M$. Therefore, in view of (E′), it follows by Lemma 1, that v^* is an integral direct sum and that $R_v/M_v=R_{v^*}/M_{v^*}$ is finitely generated over $R_{\bar{v}}/M_{\bar{v}}$ and hence over R/M. Again, in view of (C′), it follows by Proposition 2 that w is an integral direct sum. Since w and v^* are both integral direct sums, we finally conclude that v is also an integral direct sum.

Thus the induction is complete and the theorem has been established.

DEFINITION 2. *In the notation of Theorem* 1, *if v is of R-dimension $n-1$ then we shall say that v is a prime divisor for (R,M).*

Remark 2. In the notation of Theorem 1, let $n=2$. Then we get a classification of valuations v of K centered in R which is parallel to the one given in Remark 1 for the case of absolute surfaces. Namely, v is of one of the following four types: (1) v is a prime divisor for (R,M), i.e., $d=1$ and v is real discrete. (2) $d=0$ and $\rho=2$; in this case, necessarily, $r=\rho=2$, i.e., v is discrete. (3) $d=0$ and $\rho=r=1$. (4) $d=0$, $\rho=1$ and $r=2$; in this case v is necessarily an integral direct sum. In cases (1), (2), and (4), R_v/M_v is necessarily finitely generated over R/M. Also note that if v is of positive R-dimension, then v is necessarily of type (1), i.e., v is a prime divisor.

2. Quadratic transformations.

LEMMA 7. *Let $R_1 \subset R_2 \subset \cdots$ be a sequence of normal integral domains with a common quotient field K such that R_i contains a unique maximal ideal M_i and $M_{i+1} \cap R_i = M_i$ for $i=1,2,\cdots$. Assume that $\bigcup_{i=1}^{\infty} R_i$ is not the valuation ring of a valuation of K. Then there exist infinitely many valua-*

tions w of K which have center M_i in R_i and for which R_w/M_w is of positive transcendence degree over R_i/M_i for each i.

Proof. The proof is essentially the same as the one given by Zariski in the special case when the (R_i, M_i) are quotient rings of corresponding subvarieties under a sequence of birational transformations, see the theorem on page 25 of [14]. To outline the main idea of the proof, first observe that for an arbitrary commutative ring A with identity the following two statements are equivalent: (1) A contains a unique maximal ideal. (2) The set of nonunits in A is an ideal.

Now let $R = \bigcup_{i=1}^{\infty} R_i$ and $M = \bigcup_{i=1}^{\infty} M_i$. We may canonically assume that $R_1/M_1 \subset R_2/M_2 \subset \cdots$. Let $D = \bigcup_{i=1}^{\infty} R_i/M_i$. Then D is a field and $R/M = D$, i.e., M is a maximal ideal in the domain R. Also observe that R is normal. Since R is not a valuation ring, there exists $x \varepsilon K$ with $x \not\varepsilon R$ and $(1/x) \not\varepsilon R$. The canonical homomorphism h of R onto D can be uniquely extended to a homomorphism H of $R[x]$ onto $D[x]$ for which $H(x) = X$, where X is a transcendental over D; see pp. 26-27 of [14]. Let p be any one of the infinitely many prime ideals in $D[X]$. Let $P = H^{-1}(p)$. By the theorem of existence of valuations, there exists a valuation w of K having center P in $R[x]$. Since $R_w/M_w \supset D(x)$, it is clear that w has the required properties. The infinitely many choices of p gives us infinitely many w of the required type.

LEMMA 8. *Let (R, M) be a two dimensional normal local domain with quotient field K. Let P be a minimal prime ideal in R. Then:* (1) *R_P is the valuation ring of a real discrete valuation w of K;* (2) *there exists at least one and at most a finite number of valuations v of K having center M in R which are composed with w, i.e., for which $R_v \subset R_w$;* and (3) *each such valuation v is discrete of rank two and R_v/M_v is a finite algebraic extension of R/M (hence in particular v is of R-dimension zero).*

Proof. (1) is proved on page 103 of [5]. The proof of (2) is the same as in Lemma 11 of [2] and is as follows: $R/(R \cap M_w)$ is a local domain of dimension one with quotient field R_w/M_w. Hence the integral closure of $R/(R \cap M_w)$ in R_w/M_w is a Dedekind domain D with a finite number of prime ideals $P_1, P_2, \cdots, P_h$ (§ 39 of [5]). Let v_i^* be the real discrete valuation of R_w/M_w with $R_{v_i^*} = D_{P_i}$. Let v_i be the valuation of R_w/M_w which is composed of w and v_i^*. Then $v_1, v_2, \cdots, v_h$ are exactly the valuations described in (2). Finally, (3) follows at once from Theorem 1.

LEMMA 9. *Let (R, M) be a two dimensional regular local domain with quotient field K. Then: (1) R is a unique factorization domain, i.e., equivalently, every minimal prime ideal in R is principal; and (2) there exist infinitely many valuations of K having center M in R which have R-dimension zero and which are discrete of rank two.*

Proof. (1) follows by Satz 8 and Satz 9 of Krull [7].[4] To prove (2) it is enough to show, in view of Lemma 8, that there exist infinitely many relatively prime irreducible nonunits in R. To show this, let x, y be a minimal basis of M. Let $P = xR$, $\bar{R} = R/P$, $\bar{M} = M/P$, $\bar{K} = R_P/(PR_P)$, $\bar{y} =$ the residue class modulo P containing y, and $w =$ the real discrete valuation of K with $R_w = R_P$. Then $\bar{M} = \bar{y}\bar{R}$ and hence $\bar{R}$ is a regular one dimensional local domain, i.e., $\bar{R} = R_{\bar{v}}$ where v is a real discrete valuation of $\bar{K}$ with $\bar{v}(\bar{y}) = 1$. Let v be the valuation of K which is composed of w and $\bar{v}$, and let us write the elements of the value group of v as lexicographically ordered pairs of integers. Let $x_m = x + y^m$ where m is a positive integer. Since (x_m, y) is a basis of M, x_m is an irreducible nonunit. Since $v(X_m) = (0, \bar{v}(\bar{y}^m)) = (0, m)$, we have that $v(x_m) \neq v(x_n)$ whenever $m \neq n$. Therefore $x_1, x_2, x_3, \cdots$ are infinitely many pairwise relatively prime irreducible nonunits in R.

Now let us recall the notion of a quadratic transform.

LEMMA 10. *Let (R, M) be a regular local domain of dimension $s > 1$ and let $x_1, x_2, \cdots, x_s$ be a minimal basis of M. Let v be a valuation of the quotient field K of R having center M in R. Suppose we have arranged the x_i so that $v(x_1) \leqq v(x_i)$ for $i = 1, 2, \cdots, s$. Let $A = R[x_2/x_1, x_3/x_1, \cdots, x_s/x_1]$, $P = A \cap M_v$, $S = A_P$ and $N = PS$. Then (S, N) is a regular local domain of dimension $t \leqq s$, v has center N in S and if by d and d^* we denote respectively the R-dimension and the S-dimension of v then we have: $s - t = d - d^*$.*

Proof. By Lemma 3 of [1], $A/(x_1A)$ can be canonically identified with a polynomial ring $A^* = (R/M)(X_2, X_3, \cdots, X_s)$ in $s - 1$ variables over R/M. Let $P^* = P/(x_1A)$. Then $A^*_{P^*}$ is a regular local ring of dimension $h \leqq s - 1$. Fix $y_2^*, y_3^*, \cdots, y_{h+1}^*$ in A^* such that

$$P^*A^*_{P^*} = (y_2^*, y_3^*, \cdots, y_{h+1}^*)A^*_{P^*}.$$

[4] We take this opportunity in pointing out that there is obviously a misprint in the statement of satz 9 of Krull [7], namely that the term "Stellenring" should be replaced by "p-reihenring." For in the proof it is used that the Anfang forms are binary which is true only for p-reihenrings; and further, the theorem is certainly not true for arbitrary two dimensional complete local rings.

Fix y_i in A belonging to the residue class y_i^*. Then it is clear that $(x_1, y_2, y_3, \cdots, y_{h+1})S = N$ and that

$$t = \dim S = \operatorname{rank} N = 1 + \operatorname{rank} P^* = 1 + h.$$

Therefore S is regular. Also $t \leqq 1 + h \leqq (s-1) + 1 = s$. Since $N \cap R = M$, we may canonically assume that $R/M \subset S/N \subset R_v/M_v$. Since the transcendence degree of $S/N = A^*/P^*$ over R/M is $s - 1 - h = s - t$, we conclude that: $s - t =$ (R-dimension of v) — (S-dimension of v) $= d - d^*$.

DEFINITION 3. *In the notation of the above lemma, S is called "the first (or immediate) quadratic transform of R along v." Let now $R_0 = R$ and let R_i be the first quadratic transform of R_{i-1} along v assuming that* $\dim R_{i-1} > 1$. *If* $\dim R_i > 1$, *for $i = 1, 2, \cdots, n-1$ then R_n will be defined and we shall say that "R_n is the n-th quadratic transform of R along v." If S is the n-th quadratic transform of R along v for some n, we shall say that "S is a quadratic transform of R along v." Finally, if S is a quadratic transform of R along some valuation v having center M in R, we shall say that "S is a quadratic transform of R."*

DEFINITION 4. *Let (R, M) and (S, N) be local domains with a common quotient field K. If $S \supset R$ and $N \cap R = M$ then we shall say that (S, N) has center M in R. If (S, N) has center M in R and if there exists a finite set of elements $x_1, x_2, \cdots, x_n$ in S such that $S = A_P$ and $N = PS$ where $A = R[x_1, x_2, \cdots, x_n]$ and $P = A \cap N$ then we shall say that "(S, N) is a finite transform of (R, M)."*

LEMMA 11. *Let (R, M) and (S, N) be local domains with a common quotient field K. Then we have the following:* (1) *If S is a finite transform of R and if R is absolute (respectively, algebraic) then S is absolute (respectively, alegbraic).* (2) *If R is regular and if S is a quadratic transform of R then S is a finite transform of R.*

Proof. Assume that S is a finite transform of R and fix $x_1, x_2, \cdots, x_n$ in S such that $S = A_P$ and $N = PS$ where $A = R[x_1, x_2, \cdots, x_n]$ and $P = A \cap N$. Let $k =$ the ground field or $k = Z$ according as R is algebraic or absolute respectively. Then we can find a finite number of elements $x_{n+1}, x_{n+2}, \cdots, x_m$ in R such that $R = A^*_{P^*}$ where $A^* = k[x_{n+1}, x_{n+2}, \cdots, x_m]$ and $P^* = A^* \cap M$. Let $A' = k[x_1, x_2, \cdots, x_m]$ and $P' = A' \cap N$. Then it is easily seen that $S = A'_{P'}$ and $N = P'S$ and hence S is absolute or algebraic according as R is absolute or algebraic respectively. The proof of (2) is entirely similar.

PROPOSITION 3. *Let (R, M) be an n-dimensional regular local domain with quotient field K with $n > 1$, and let v be a valuation of K which is a prime divisor for R. Then the quadratic sequence along v starting from R is necessarily finite, i.e., if $R = R_0$ and if R_i is the first quadratic transform of R_{i-1} along v provided $\dim R_{i-1} > 1$ then for some integer h we have that R_h is one dimensional; we also have: $R_h = R_v$. Furthermore, there exists a field T with $R/M \subset T \subset R_v/M_v$ such that T is finitely generated over R/M and R_v/M_v is a pure transcendental extension of T of finite positive transcendence degree (we may express this by saying that "R_v/M_v is a ruled extension of R/M").*

Proof. Assume the contrary, i.e., that the quadratic sequence $R = R_0 < R_1 < R_2 < \cdots$ along v is infinite. It then follows, by Lemma 10, that there exists an integer s such that $\dim R_t = \dim R_s$ and R_t-dimension of $v = m - 1$ whenever $t \geqq s$, where we have set $m = \dim R_s$. Let $S = \bigcup_{i=0}^{\infty} R_i$ and $N = \bigcup_{i=0}^{\infty} M_i$ where M_i is the maximal ideal in R_i. Then, as in the proof of Lemma 7, N is the unique maximal ideal in S and $\bigcup_{i=s}^{\infty} R_i/M_i = S/N$. Since, as in the proof of Lemma 10, R_{t+1}/M_{t+1} is an algebraic extension of R_t/M_t whenever $t \geqq s$, it follows that S/N is an algebraic extension of R_t/M_t for any $t \geqq s$. Now v has center M_i in R_i for each i and by Theorem 1, v is real discrete. Hence, if S were not the valuation ring of a valuation of K then, following the considerations of pages 27-28 of [14], we would reach a contradiction. Therefore, S must be the valuation ring of a valuation w of K. Since R_s has a first quadratic transform R_{s+1}, it follows by the definition of quadratic transforms that $m > 1$, i.e., that v is of positive R_t-dimension whenever $t \geqq s$. Now $R_v \supset R_w$ and

$$R_w \cap M_v = S \cap M_v = \left(\bigcup_{i=0}^{\infty} R_i\right) \cap M_v = \left(\bigcup_{i=0}^{\infty} (R_i \cap M_v)\right) = \bigcup_{i=0}^{\infty} M_i = N = M_w.$$

Therefore $v = w$, i.e., $R_v/M_v = S/N$ is an algebraic extension of R_s/M_s. Thus our assumption that the quadratic sequence $R_0 < R_1 < R_2 < \cdots$ is infinite is absurd. Let R_h be the last member of this sequence. Then R_h is a regular one dimensional local domain, i.e., R_h is the valuation ring of a real discrete valuation of K. Since $R_h \subset R_v$, we must have $R_h = R_v$. Now let $T = R_{h-1}/M_{h-1}$ and let d be the dimension of R_{h-1}. Then $d > 1$. Let $x_1, x_2, \cdots, x_d$ be a minimal basis of M_{h-1} arranged so that $v(x_1) \leqq v(x_i)$ for $i = 1, 2, \cdots, d$. Let $A = R_{h-1}[x_2/x_1, x_3/x_1, \cdots, x_d/x_1]$ and $P = A \cap M_h$. As in Lemma 3 of [1], we may identify $A/(x_1A)$ with the polynomial ring

$A^* = T[Y_2, Y_3, \cdots, Y_d]$. Let $P^* = P/(x_1 A)$. As in the proof of Lemma 10, $\operatorname{rank} P^* = (\dim R_h) - 1 = 0$, i.e., $P^* = (0)$. Since $R_v/M_v = R_h/M_h$ is isomorphic to A^*/P^*, i.e., to $T(Y_2, Y_3, \cdots, Y_d)$, we conclude that $R_v M/_v$ is a pure transcendental extension of T of transcendence degree $d - 1 > 0$. By Theorem 1, R_v/M_v is finitely generated over R/M and hence T is also finitely generated over R/M; this also follows from the fact that R_{h-1} is a finite transform of R.

PROPOSITION 4. *Let V be an r-dimensional algebraic variety with function field K/k, let v be a prime divisor of K/k, and let W be the center of v on V. Let s be the dimension of W. Assume that v is of the second kind for V (i.e., that $s < r - 1$) and that W is simple for V. Let V^* be any other projective model of K/k such that V is of the first kind for V^* and such that V^* is normal at the center of W^* of v on V^*. Let L/k be the function field of W^*. Then we have the following:* (1) *W^* is a ruled variety, i.e., we can find a field T with $k \subset T \subset L$ such that L is a pure transcendental extension of T of positive transcendent degree.* (2) *If $r = 2$, then we can find a field T with $k \subset T \subset L$ such that T/k is a finite algebraic extension and L/T is a pure transcendental extension of transcendence degree one.* (3) *If $r = 2$ and if k is algebraically closed, then L/k is a pure transcendental extension of transcendence degree one, i.e., W^* is a rational curve.*

Proof. (1) follows from Proposition 3 by observing that $R_v/M_v = L$ and that if (R, M) denotes the quotient ring of W on V then $k \subset R/M \subset R_v/M_v$. Again, (2) follows from (1) and (3) follows from (2).

LEMMA 12. *Let $R_0 < R_1 < R_2 < \cdots$ be a strictly ascending sequence of two dimensional regular local domains with a common quotient field K, let M_i be the maximal ideal in R_i, and let $S = \bigcup_{i=0}^{\infty} R_i$. Assume that R_{i+1} is a quadratic transform of R_i for $i = 0, 1, 2, \cdots$. Then:* (1) *S is the valuation ring of a valuation v^* of K such that v^* has center M_i in R_i and such that v^* is of R_i-dimension zero for each i. Furthermore:* (2) *if v is any valuation of K with center M_i in R_i for $i = 0, 1, 2, \cdots$, then $v = v^*$.*

Proof. By introducing all the intermediate successive quadratic transforms between R_i and R_{i+1} for each i, we may assume without loss of generality, since this would not change S, that R_{i+1} is a first quadratic transform of R_i for $i = 0, 1, 2, \cdots$. Now suppose, if possible, that S is not the valuation ring of a valuation of K. Then by Lemma 7, there exists a valuation w of K having center M_i in R_i and of positive R_i-dimension for each i; and hence

by Theorem 1, w is a prime divisor for R_0. Therefore by Proposition 3, the sequence $R_0 < R_1 < R_2 < \cdots$ must be finite, which is absurd. Hence $S = R_{v^*}$ where v^* is a valuation of K. Since, as in the proof of Lemma 7, $M_{v^*} = \bigcup_{i=0}^{\infty} M_i$, it follows that

$$R_i \cap M_{v^*} = R_i \cap (\bigcup_{j=0}^{\infty} M_j) = R_i \cap (\bigcup_{j=i+1}^{\infty} M_j) = \bigcup_{j=i+1}^{\infty} (R_i \cap M_j) = M_i,$$

i.e., v^* has center M_i in R_i for each i. Again by Proposition 3 and Theorem 1, it follows that the R_i-dimension of v^* is zero for each i. Finally, if v is any other valuation of K with center M_i in R_i for each i then $R_v \supset \bigcup_{i=0}^{\infty} R_i = R_{v^*}$ and $M_v \cap R_{v^*} = \bigcup_{i=0}^{\infty} (M_v \cap R_i) = \bigcup_{i=0}^{\infty} M_i = M_{v^*}$ and hence $v = v^*$.

LEMMA 13. *Let (R, M) be a regular two dimensional local domain with quotient field K, let P be a minimal prime ideal in R, let $\bar{R} = R/P$ and $\bar{M} = M/P$. Let $\bar{K}$ be the quotient field of $\bar{R}$ and let T be the integral closure of $\bar{R}$ in $\bar{K}$. Assume that R is the quotient ring of a point either on an algebraic surface or on an absolute surface (i.e., K is a one dimensional algebraic function field over Q and R is the quotient ring of a finitely generated domain over Z with respect to a maximal ideal). Then T is a finite R-module.*

Proof. The case of algebraic surfaces is well known (see p. 511 of [10]) and so we may assume that R is the quotient ring of a point on an absolute surface. Then we can find $z_1, z_2, \cdots, z_s$ in K such that K is the quotient field of $R^* = Z[z_1, z_2, \cdots, z_s]$ and such that $R = R^*_{M^*}$ where M^* is a maximal ideal in R^*. Let $P^* = R^* \cap P$, $\bar{R}^* = R^*/P^*$, $\bar{M}^* = M^*/P^*$, and $\bar{z}_i =$ the P^*-residue class of z. Then $\bar{R} = \bar{R}^*_{\bar{M}^*}$. If $Z \cap P^* \neq (0)$ then $Z \cap P^* = pZ$ where p is a prime number and $\bar{R}^* = (Z/pZ)[\bar{z}_1, z_2, \cdots, \bar{z}_s]$; hence we are again in the algebro-geometric case. Now assume that $Z \cap P^* = (0)$. Let $S = R^*_{P^*}$, $N = P^*S$ and $\bar{S} = S/N$. Then $\bar{S} = \bar{K}$. Also $Q \subset S$ and hence $S = S^*_{N^*}$ where $S^* = Q[z_1, z_2, \cdots, z_s]$ and $N^* = S^* \cap N$. Since S^* is of transcendence degree one over Q, it follows that $\bar{K}$ is a finite algebraic extension of Q. Thus it is enough to prove *the following assertion: Let L be a finite algebraic extension of Q, let A be a proper subdomain of L having L as quotient field, and let B be the integral closure of A in L. Then B is a finite A-module.*

First let us recall that if D is a commutative ring with 1, and W the set of all maximal ideal in D then $D = \bigcap_{w \in W} D_w$. For if $d \in \bigcap_{w \in W} D_w$ then $d = e_w/f_w$ with $e_w, f_w \in D$ and $f_w \notin w$. Now since the ideal E generated by all the f_w

is not contained in any member of W, we must have $E = D$. Therefore we can write $1 = f_{w_1}g_1 + f_{w_2}g_2 + \cdots + f_{w_h}g_h$ with g_i in D and hence $d = e_{w_1}g_1 + e_{w_2}g_2 + \cdots + e_{w_h}g_h \varepsilon D$. Therefore $D = \bigcap_{w \varepsilon W} D_w$. To prove the italicized assertion, let Y be the integral closure of Z in L. Then Y is a finite Z-module; let $y_1, y_2, \cdots, y_n$ be a module basis of Y over Z. Let $C = A[y_1, y_2, \cdots, y_n]$. Then it is clear that $y_1, y_2, \cdots, y_n$ is a module basis of C over A. Let M be a maximum ideal in C, and let $m = Y \cap M$. Then $Y_m \subset C_M$. If $m = (0)$ then $Y_m = L$ and if $m \neq (0)$ then Y_m is the valuation ring of real discrete valuation of L, i.e., Y_m is a maximal subring of L. Therefore $Y_m = C_M$. Hence by the above observation, C is an intersection of valuation rings. Therefore C is integrally closed in L. Since $C \subset B$, we must have $C = B$, i.e., B is a finite A-module.

In the following proposition we shall prove that the singularities of a curve lying on a nonsingular *absolute surface* can be resolved by quadratic transformations applied to the surface. We shall base our demonstration on Zariski's proof of the algebro-geometric case; see Theorem 4 of [11].

PROPOSITION 5. *Let (R, M) be a two dimensional regular quotient ring of a point on an algebraic or absolute surface and let K be the quotient field of R. Let P be a minimal prime ideal in R and let w be the valuation of K with $R_w = R_P$. Let v be a valuation of K composite with w and having center M in R. Let (R_n, M_n) be the n-th quadratic transform of R along v and let $P_i = R_i \cap M_v$. Then P_i is a minimal prime ideal in R_i for $i = 1, 2, \cdots$; and there exists an integer n such that for any $i \geqq n$ we can choose a basis (x_i, y_i) of M_i for which $x_iR_i = P_i$, $v(y_i) = (0,1)$, and $v(x) = (1, a)$ where a is some integer.*

Proof. Since $M_i \cap R = M$ and $P_i \cap R = R \cap M_w = P$ and since R_i is two dimensional, P_i must be a minimal prime ideal in R_i. Now let (x, y) be a basis of M, and suppose for instance that $v(x) \leqq v(y)$. Then $w(x) \leqq w(y)$. Therefore $x \not\varepsilon P$; for otherwise we would have $w(x) > 0$ and hence $w(y) > 0$, i.e., $x \varepsilon R \cap M_w = P$ and $y \varepsilon R \cap M_w = P$, and hence $M = P$, which is a contradiction. Let $R^* = R[y/x]$, $M^* = R^* \cap M_v$, $P^* = R^* \cap M_w$. Let $\bar{R} = R/P$, $\bar{M} = M/P$, $\bar{R}^* = R^*/P^*$, $\bar{M}^* = M^*/P^*$, $\bar{R}_i = R_i/P_i$, $\bar{M}_i = M_i/P_i$, $\bar{K} = R_w/M_w$, $\bar{v} =$ the valuation of $\bar{K}$ induced by v, and $\bar{x}, \bar{y}$ the residue classes modulo M_w respectively of x, y. Then $(\bar{R}_i, \bar{M}_i)$ is a one dimensional local domain with quotient field $\bar{K}$, the real discrete valuation $\bar{v}$ of $\bar{K}$ has center $\bar{M}_i$ in $\bar{R}_i$, and $\bar{R} = \bar{R}_0 \subset \bar{R}_1 \subset \bar{R}_2 \subset \cdots$. Now $R_1 = R^*_{M^*}$, $M_1 = M_0R_1$ and $0 \neq \bar{x} \varepsilon \bar{R}$. Therefore $(\bar{x}, \bar{y})\bar{R} = \bar{M}$, $\bar{R}^* = \bar{R}[\bar{y}/\bar{x}]$, $\bar{M}^* = \bar{R}^* \cap M_{\bar{v}}$ and

$\bar{R}_1 = \bar{R}^*_{\bar{M}^*}$. Hence $\bar{M}\bar{R}^* = \bar{x}\bar{R}^*$ and $\bar{M}\bar{R}_1 = \bar{x}_1\bar{R}_1$. Similarly $\bar{M}_i\bar{R}_{i+1} = z_{i+1}\bar{R}_{i+1}$ with $z_{i+1} \varepsilon \bar{R}_{i+1}$ for $i = 0, 1, 2, \cdots$.

We shall now show that $\bigcup_{i=0}^{\infty} \bar{R}_i = R_{\bar{v}}$. Given c in $R_{\bar{v}}$ we can write $c = a/b$ with $0 \neq b$, $a \varepsilon \bar{R}$. If $b \not\varepsilon \bar{M}$ then $c \varepsilon \bar{R}$. Suppose that $b \varepsilon \bar{M}$. Since $\bar{v}(a) \geqq \bar{v}(b)$ we must have $a \varepsilon \bar{M}$. Hence $b = b_1 z_1$, $a = a_1 z_1$ with a_1, b_1 in $\bar{R}_1$. Since $c = a_1/b_1$, again either $c \varepsilon \bar{R}_1$ or $a_1, b_1 \varepsilon \bar{M}_1$. Suppose that $c \not\varepsilon \bar{R}_1$. Then $a_1 = a_2 z_2$, $b_1 = b_2 z_2$ with $a_2, b_2 \varepsilon \bar{R}_2$. Similarly if $c \not\varepsilon \bar{R}_{n-1}$ then $a = a_n z_1 z_2 \cdots z_n$ and $b = b_n z_1 z_2 \cdots z_n$ with $a_i, b_i \varepsilon \bar{R}_i$. Since $\bar{M}_{n-1} = \bar{R}_{n-1} \cap M_{\bar{v}}$, we must have $\bar{v}(z_n) > 0$, and hence that $\bar{v}(b) \geqq n$ where we take Z as the value group of the real discrete valuation $\bar{v}$. Since $\bar{v}$ is real discrete and since $b \neq 0$, for some n we must have $c \varepsilon \bar{R}_n$. Therefore $\bigcup_{i=0}^{\infty} \bar{R}_i = R_{\bar{v}}$. Let S_i be the integral closure of $\bar{R}_i$ in $\bar{K}$. Then S_i is a Dedekind domain with a finite number of prime ideals (§ 39 of [5]), and since the quotient ring of S_i with respect to any prime ideal is a real discrete valuation ring, we must have $S_i = \bigcap_{u \varepsilon W_i} R_u$ where W_i is a finite set of real discrete valuations of $\bar{K}$. It is clear that the valuations in W_i are exactly the valuations of $\bar{K}$ having center $\bar{M}_i$ in $\bar{R}_i$. Therefore $W_i \supset W_{i+1}$. Let $u_1, u_2, \cdots, u_h$ be the valuations in W_0 different from $\bar{v}$. Since $R_{\bar{v}}$ is a maximal subring of $\bar{K}$, we can find $a_i \varepsilon R_{\bar{v}}$ such that $a_i \not\varepsilon R_{v_i}$. Since $\bigcup_{j=0}^{\infty} \bar{R}_j = \bar{R}_{\bar{v}}$, $a_i \varepsilon \bar{R}_{m_i}$ for some integer m_i. Let $m = \max(m_1, m_2, \cdots, m_h)$. Then $a_i \not\varepsilon \bar{R}_m$ for $i = 1, 2, \cdots, h$. Therefore $W_m = \{\bar{v}\}$, i.e., $R_{\bar{v}}$ is the integral closure of $\bar{R}_m$. By Lemma 13, $R_{\bar{v}}$ is a finite module over $\bar{R}_m$. Hence by the Hilbert basis theorem, we can find n such that $\bar{R}_i = \bar{R}_{\bar{v}}$ whenever $i \geqq n$; from now on we shall assume that $i \geqq n$. Fix $\bar{y}_i$ in $\bar{R}_i$ with $\bar{v}(\bar{y}_i) = 1$. Let y_i be an element of R_i belonging to the residue class $\bar{y}_i$. Then $v(y_i) = (0, 1)$. By Lemma 9, P_i is a principal ideal; let $P_i = x_i R_i$. Since $(R_i)_{P_i} = R_w$, $v(x_i) = (1, a)$. Finally given z in M_i let $\bar{z}$ be the residue class of z modulo P_i. Then $\bar{z} \varepsilon \bar{M}_i$ and hence $\bar{z} = \bar{y}_i \bar{t}$ with $\bar{t} \varepsilon \bar{R}_i$. Fix t in R_i belonging to the residue class $\bar{t}$. Then $z - y_i t \varepsilon P_i$, i.e., $z \varepsilon (x_i, y_i) R_i$. Therefore $(x_i, y_i) R_i = M_i$.

Remark 3. Observe that Lemma 13, which is crucial in the proof of Proposition 5, breaks down for certain abstract local rings (regular of dimension two) as can be shown by using an example due to F. K. Schmidt; see page 24 of Zariski's paper [12]. To indicate this we shall use Zariski's notation. We have then two independent variables x and t over a certain field k of characteristic $p \neq 0$, $\Sigma' = k(x, t)$, o' is the valuation ring of a certain real discrete valuation v' of Σ'/k, $\Sigma = \Sigma'(\tau)$ with $\tau = t^{1/p}$ and o is the

one dimensional local domain $o'[\tau]$ with quotient field Σ. Let T be the integral closure of o' in Σ. Since o is integral over o', T is also the integral closure of o in Σ. Let $R = o'[X]$ and $P = (X^p - t)R$. Then it is easily verified that R is a two dimensional regular local domain, P is a minimal prime ideal in R, and $X \rightarrow \tau$ is an o'-homomorphism of R onto o with kernel P, i.e., o is isomorphic to R/P. Since T is not a finite o'-module (see part c on page 447 of [9]) and since o is a finite o'-module, T cannot be a finite o-module. Thus Lemma 13 is not applicable to the minimal prime ideal P of the two dimensional regular local domain R.

Now let K be the quotient field of R, let M be the maximal ideal in R, let w be the valuation of K with $R_w = R_P$, and let v be a valuation of K composed with w and having center M in R. Suppose, if possible, that Proposition 5 were true for the regular local domain (R, M). Then, in the notation of Proposition 5, we would have that R_n/P_n is a valuation ring. From the considerations made in the proof of Proposition 5, it follows that this would imply that o can be transformed into a regular local ring by applying to o, n successive quadratic transformations in the sense described on page 24 of [12], but this is impossible as is proved on pages 24-25 of [12]. Therefore, Proposition 5 is not applicable to the valuation v having center M in the regular two dimensional local domain R. Since Proposition 5 is essentially based on Lemma 13,[5] this again shows, as we have directly proved above, that Lemma 13 is not valid for arbitrary two dimensional regular local domains.

THEOREM 2.[6] *Let (R, M) be a two dimensional regular local domain with quotient field K. Let v be a valuation of K having center M in R and R-dimension zero. Let f be a given nonzero element of R. Then there exists a quadratic transform (R^*, M^*) of R along v and a basis (x^*, y^*) of M^* such that $f = x^{*a}y^{*b}d$ where a and b are nonnegative integers, d is a unit in R^*, and where the following conditions are satisfied:* (A) *If v is real of rational rank one, then $b = 0$.* (B) *If v is real of rational rank two then either $b = 0$ or $v(x^*)$ and $v(y^*)$ form an integral basis for the value group of v.* (C) *If v is of rank two and if R is the quotient ring of a point either on an algebraic surface or on an absolute surface, then $v(x^*) = (1, h)$ and $v(y^*) = (0, 1)$ where we are writing the v-values of elements of K as lexicographically ordered pairs of integers and where h is some integer.*

---

[5] I. e., the only fact about valuations centered in local domains which is used in the proof of Proposition 5 and in which the local domains are qualified to be either algebraic or absolute is Lemma 13.

[6] See Proposition 3 of [2].

Proof. First assume that v is real. Let $R_0 = R$ and $M_0 = M$. Let (x_0, y_0) be a basis of M_0 and let (R_i, M_i) be the i-th quadratic transform of R_0 along v. We shall define elements (x_i, y_i) of R_i by induction on i. Let then $i = m > 0$ and assume that we have defined x_i, y_i for $i = 1, 2, \cdots, m-1$. Suppose first that $v(y_{i-1}) \geqq v(x_{i-1})$. Let $S_i = R_{i-1}[y_{i-1}/x_{i-1}]$ and $N_i = M_{i-1}S_i$. Let $P_i = M_v \cap S_i$. Let z_i be the residue class modulo N_i containing y_{i-1}/x_{i-1}. Then by Lemma 3 of [1], $S_i/N_i = (R_{i-1}/M_{i-1})[z_i]$ and z_i is transcendental over R_{i-1}/M_{i-1}. Since P_i is a maximal ideal in S_i containing N_i, P_i/N_i must be a maximal ideal in S_i/N_i and hence $P_i/N_i = g_i(z_i)(S_i/N_i)$ where $g_i(X)$ is a monic irreducible polynomial in $(R_{i-1}/M_{i-1})[X]$. Let $G_i(X)$ be a monic polynomial in $R_{i-1}[X]$ which when reduced modulo M_{i-1} gives $g_i(X)$. We set $x_i = x_{i-1}$ and $y_i = G_i(y_{i-1}/x_{i-1})$. Secondly, if $v(y_{i-1}) < v(x_{i-1})$ then we set $x_i = y_{i-1}$ and $y_i = x_{i-1}/y_{i-1}$. Then by Corollary 1 of [1] $(x_i, y_i)R_i = M_i$.

Let $f_0 = f$ and define by induction $f_i \in R_i$ by the equation $f_{i-1} = x_i^{u_i} f_i$ where u_i is a nonnegative integer and where f_i is an element in R_i prime to x_i (by Lemma 9, R_i is a unique factorization domain). Let $f_{i,1}, f_{i,2}, \cdots, f_{i,h_i}$ be the irreducible factors of f_i in R_i and let $w_{i,j}$ be the valuation of K whose valuation ring is the quotient ring of R_i with respect to $f_{i,j}R_i$. Let W_i be the set of valuations u of K such that u has center M_i in R_i and u is composed with $w_{i,j}$ for some j. By Lemma 8, W_i is a finite set. For a given u in W_i, let $P_i = M_u \cap R_{i-1}$. Since u is nontrivial, $P_i \neq (0)$. Since $x_i \not\in M_u \cap R_i$ and since $x_i R_i = M_{i-1}R_i$, $P_i \neq M_{i-1}$. Therefore P_i is a minimal prime ideal in R_{i-1}. Since R_{i-1} is a unique factorization domain and since $f_{i-1} \in P_i$, we must have $P_i = f_{i-1,j}R_{i-1}$ for some j, i.e., $u \in W_{i-1}$. Thus $W_i \subset W_{i-1}$. Since by Lemma 12, $\bigcup_{i=1}^{\infty} R_i = R_v$, since v is real, and since no element of W_i is real, it follows that $\bigcap_{i=1}^{\infty} W_i = \emptyset$. Since W_i is finite, $W_m = \emptyset$ for some m, i.e., f_m is a unit in R_m. Hence we have $f = x_m{}^A y_m{}^B D$ where A and B are nonnegative integer and D is a unit in R_m. If either v is of rational rank one or if v is of rational rank two and $v(x_m)$ and $v(y_m)$ are rationally dependent, then the proof can be completed by the argument of the last part of the proof of Proposition 3 of [2]. Now suppose that v is of rational rank two and that $v(x_m)$ and $v(y_m)$ are rationally independent. We may then take R^* to be R_m and $x^* = x_m$, $y^* = y_m$. It then remains to be shown that $v(x^*)$ and $v(y^*)$ form an integral basis for the value group of v. Let $v(x^*) = p$ and $v(y^*) = q$; and suppose for instance that $p < q$. Fix a representative system k in R^* of R^*/M^* (k is not in general a field; note that we take zero as the representative of M^*/M^*). Let z be an arbitrary nonzero element of R^* and let $v(z) = r$. Fix an integer n so that $r < np$. Then we can find

a polynomial $H(X,Y)=\sum_{i+j\leqq n} H_{ij}X^iY^j$ of degree at most n with coefficients H_{ij} in k such that $z^*=z-H(x^*,y^*)\,\varepsilon\, M^{*n}$ (see § 5 of [6]). Since $v(u)>r$ for any $u\,\varepsilon\, M^{*n}$, we must have that $v(z^*)>r$ and hence that $r=v(z)=v(H(x^*,y^*))$. Since p and q are rationally independent and since $v(H_{ij})=0$ whenever $H_{ij}\neq 0$, we can find $H_{st}\neq 0$ such that $v(H_{st}x^{*s}y^{*t})<v(H_{ij}x^{*i}y^{*j})$ whenever $H_{ij}\neq 0$ and whenever either $i\neq s$ or $j\neq t$. Therefore $r=v(z)=v(H(x^*,y^*))=v(H_{st}x^{*s}y^{*t})=sp+tq$. Therefore for any nonzero element z_1 of K, we must have $v(z_1)=s_1p+t_1q$ where s_1 and t_1 are integers. Thus we have shown that $\{p,q\}$ is an integral basis of the value group of v. This completes the proof of (A) and (B).

To prove (C), assume that v is of rank two, that R is the quotient ring of a point either on an algebraic or on an absolute surface and that we are writing the values of elements of K as lexicographically ordered pairs of integers. By Proposition 5, we can find a quadratic transform $(\bar{R},\bar{M})$ of R along v and a basis $\bar{x},\bar{y}$ of $\bar{M}$ such that $v(\bar{x})=(1,p)$ and $v(\bar{y})=(0,1)$ where p is some integer. Let $(\bar{R}_i,\bar{M}_i)$ be the i-th quadratic transform of $\bar{R}$ along v. Let $\bar{x}_i=\bar{x}/\bar{y}_i$ and $\bar{y}_i=\bar{y}$. Since $v(\bar{x})>iv(\bar{y})$ for any integer i, it follows that $(\bar{x}_i,\bar{y}_i)\bar{R}_i=\bar{M}_i$. Let $f=\bar{x}^ag_0$ where a is a nonnegative integer and where g_0 is an element of $\bar{R}$ prime to $\bar{x}$. Then $v(g_0)=(0,t)$ where t is some nonnegative integer. Define g_i by induction by the equation $g_{i-1}=\bar{y}_i^{e_i}g_i$ where e_i is a nonnegative integer and g_i is an element of $\bar{R}_i$ prime to $\bar{y}_i$. Since $M_{i-1}\bar{R}_i=\bar{y}_i\bar{R}_i$, we have that $e_i>0$ whenever g_{i-1} is a nonunit in $\bar{R}_{i-1}$ (i.e., whenever $g_{i-1}\,\varepsilon\,\bar{M}_{i-1}$). Therefore, if g_{i-1} is a nonunit in $\bar{R}_{i-1}$ for $i=1,2,\cdots,n$, then $v(g_0)\geqq(0,n)$. Hence for some integer $m\leqq t$, g_m must be a unit in $\bar{R}_m$. Hence $f=x^{*a}y^{*b}d$ where b is a nonnegative integer, d is a unit in $R^*=\bar{R}_m$, $x^*=\bar{x}_m$, and $y^*=\bar{y}_m$.

We shall now prove a generalization to abstract two dimensional regular local domains of Zariski's theorem on the factorization of birational transformation between nonsingular algebraic surfaces into local quadratic transformations; see the lemma on page 538 of [11].

THEOREM 3. *Let (R,M) and (R',M') be regular two dimensional local domains with a common quotient field K such that (R',M') has center M in R. Then R' is a quadratic transform of R.*

We shall precede the proof proper of this theorem by three preparatory lemmas.

LEMMA 14. *Let A be a normal domain with quotient field K such that A contains a unique maximal ideal P and let $0\neq x\,\varepsilon\, K$ such that $x\notin A$ and*

$(1/x) \not\in A$. Let $B = A[x]$ and let $M = PB$. Then: (1) *there exists a valuation v of K such that $R_v \supset B$, $M_v \cap B = M$;* (2) *$M \cap A = P$;* (3) *the residue class $\bar{x}$ modulo M containing x is transcendental over A/P, i.e., R_v/M_v contains the polynomial ring $(R/M)[\bar{x}]$ and hence R_v/M_v is of positive transcendence degree over R/M.*

Proof. Follows by the considerations made by Zariski on pages 26-27 of [14].

LEMMA 15. *Let (R, M) be an n dimensional regular local domain with quotient field K and $n > 1$. Let $\{x_1, x_2, \cdots, x_n\}$ be a minimal basis of M. Let $A = R[x_2/x_1, x_3/x_1, \cdots, x_n/x_1]$ and $P = MA$. Then:* (1) *$P = x_1A$;* (2) *P is a minimal prime ideal of A;* (3) *A_P is the valuation ring of a real discrete valuation w of K;* (4) *$A/P = (R/M)[\bar{y}_2, \bar{y}_3, \cdots, \bar{y}_n]$ where $\bar{y}_i =$ the residue class modulo P containing x_i/x_1, and $\bar{y}_2, \bar{y}_3, \cdots, \bar{y}_n$ are algebraically independent over R/M;* (5) *$R_w/M_w = (R/M)(\bar{y}_2, \bar{y}_3, \cdots, \bar{y}_n)$; and* (6) *$w$ is the "M-adic divisor of R," i.e., w is completely defined by setting $w(a/b) =$ (leading degree of a) — (leading degree of b) where a and b are any two nonzero elements of R, and w is of R-dimension $(n-1)$.*

Proof. (1) is obvious, (2) and (4) are proved in Lemma 3 of [1]. Now let N be a maximal ideal in A containing P, let $R^* = A_N$ and $M^* = NR^*$. Then by Corollary 1 of Lemma 3 of [1], R^* is a regular local domain and $P^* = y_1R^*$ is a minimal prime ideal of R^*. Hence $R^*_{P^*} = R_w$ where w is a real discrete valuation of K. Now (3) follows since $A_P = R^*_{P^*}$, and (5) follows from (4) and the fact that $A_P/(PA_P)$ is the quotient field of A/P. Now let a be an arbitrary nonzero element of R and let d be the leading degree of a. Then $a = f(x_1, x_2, \cdots, x_n)$ where $f(X_1, X_2, \cdots, X_n)$ is a form of degree d with coefficients in R not all in M; and hence $a = x_1{}^d g$ where $g = f(1, x_2/x_1, x_3/x_1, \cdots, x_n/x_1) \not\in P$ since $\bar{y}_2, \bar{y}_3, \cdots, \bar{y}_n$ are algebraically independent over R/M; i.e., $w(a) = d$. Finally, it follows by (5) that w is of R-dimension $(n-1)$ and this proves (6).

LEMMA 16. *Let (R, M) and (R', M') be as in Theorem 3 and assume that $R \neq R'$. Then there exists a nonunit z in R' such that $MR' \subset zR'$.*

Proof. Suppose, if possible, that MR' is primary for M'. Let $(\bar{R}, \bar{M})$ and $(\bar{R}', \bar{M}')$ be the completions of R and R' respectively. Let w be the M'-adic divisor of R'. Then w has center M in R and R-dimension of $w \geqq R'$-dimension of $w = 1$. Therefore by Theorem 1, w is a prime divisor also for R. Hence by Theorem 2 of [13], (R, M) and (R', M') are both subspaces

of (R_w, M_w) and hence R is a subspace of R'. Therefore we may canonically assume that $\bar{R}$ is a subring and a subspace of $\bar{R}'$. By Theorem 1, R_w/M_w is finitely generated over R/M and hence R'/M' must be a finite algebraic extension of R/M. Since $\bar{R}'/\bar{M}' = R'/M'$ and $\bar{R}/\bar{M} = R/M$, we have that $\bar{R}/\bar{M}$ is a finite algebraic extension of R/M. Also MR' is primary for M' implies that $\bar{M}'\bar{R}'$ is primary for $\bar{M}'$. Therefore by Theorem 8 of [3], $\bar{R}'$ is integral over $\bar{R}$. Let E and E' be the quotient fields of $\bar{R}$ and $\bar{R}'$ respectively. Let A be the set of valuations v of K having center M in R and let B be the set of valuations u of K for which $R_u \supset R$. Let v be a member of A. Then by Lemma 13 of [2], v has an extension $\bar{v}$ to E having center $\bar{M}$ in $\bar{R}$. Since $\bar{R}'$ is integral over $\bar{R}$, we must have $R_{\bar{v}'} \supset \bar{R}'$ for some extension $\bar{v}'$ of $\bar{v}$ to E'. Since $R_v = R_{\bar{v}'} \cap K$ and since $R' = \bar{R}' \cap K$ (Lemma 2 of [2]), we must have $R' \subset R_v$. Therefore $R' \subset \bigcap_{v \varepsilon A} R_v$. Now let u be a member of B for which $P = M_u \cap R \neq M$. Let $\bar{K} = R_u/M_u$, $\bar{R} = R/P$ and $\bar{M} = M/P$. Let $\bar{u}$ be a valuation of $\bar{K}$ having center $\bar{M}$ in $\bar{R}$ and let u^* be the valuation of K which is composed of u and $\bar{u}$. Then $R_{u^*} \supset R_u$ and u^* has center M in R. Therefore $R' \subset \bigcap_{u \varepsilon B} R_u = R$, i.e., $R' = R$. Thus our assumption that MR' is primary for M' is absurd. Since R' is two dimensional, $MR' \subset p$ where p is a minimal prime ideal in R' and by Lemma 9, $p = zR'$ with $z \varepsilon R'$, i.e., $MR' \subset zR'$.

Proof of Theorem 3. If $R = R'$ then there is nothing to prove. So assume that $R < R'$. By Lemma 9, there exists a discrete rank two valuation v of K having center M' in R'. By Theorem 1, R-dimension of $v = 0 = R'$-dimension of v. Let $\{x, y\}$ be a minimal basis of M, and suppose for instance that $v(x) \leqq v(y)$. Let $t = y/x$. If $1/t \varepsilon R'$ then $t \varepsilon R_v$ and $1/t \varepsilon R_v$ and hence $1/t \not\varepsilon M_v$, i.e., $1/t \not\varepsilon M'$ and hence $t \varepsilon R'$. Now suppose, if possible, that $t \not\varepsilon R'$. Then $1/t \not\varepsilon R'$. Let $A' = R'[t]$ and $P' = M'A'$. By Lemma 14, there exists a valuation w of K with center P in A, and

$$R_w/M_w \supset A'/P' = (R'/M')[\bar{t}]$$

where $\bar{t}$ is the residue class modulo P' containing t and $\bar{t}$ is transcendental over R'/M'. Let $A = R[t]$, $P = MA$ and $P^* = P' \cap A$. Then $P^* \cap R = M$ and hence t is *a fortiori* transcendental over R modulo M. Hence $P^* = MA = P = xA$. Therefore $A_P \supset R_w$ and hence by Lemma 15, w must be the M-adic divisor of R. Since $w(x) = 1$ and since $w(M') > 0$, we have that $x \varepsilon M'$ and $x \not\varepsilon (M')^2$, i.e., that x is an irreducible nonunit in R'. By Lemma 16, $x = az$ and $y = bz$ where a and b are in R' and z is a nonunit in R'. Since by Lemma 9, R' is a unique factorization domain, we must have $y = cx$

with c in R', i.e., $t = c \varepsilon R'$. Thus our assumption that $t \notin R'$ is absurd. Let (R_i, M_i) be the i-th quadratic transform of R along v. Then $R_1 \subset R'$ and R' has center M_1 in R_1. Similarly, if $R_1 < R'$ then $R_2 \subset R'$ and R' has center M_2 in R_2 and so on. Suppose, if possible, that $R_i \subset R'$ for all i. Then by Lemma 12, $R_v \subset R'$ and hence $R_v = R'$. Since v is rank two, R_v, i.e., R' is nonnoetherian which is absurd. Therefore, for some integer n, we must have $R_n = R'$.

LEMMA 17. *Lemma 12 remains true if we replace the assumption that R_{i+1} is a quadratic transform of R_i for $i = 0, 1, 2, \cdots$, by the weaker assumption that $M_{i+1} \cap R_i = M_i$ for $i = 0, 1, 2, \cdots$.*

Proof. Follows by Lemma 12 and Theorem 3.

Appendix.*

That an ordered abelian group G of finite rank can be embeded (as an ordered subgroup) in the lexicographically ordered direct sum $R^{(\rho)}$ of ρ copies of the additive group R of real numbers can be proved as follows: First we prove that *if G is rationally complete* (i.e., with any element g of G and an arbitrary nonzero integer n, G always contains an element g^* such that $ng^* = g$) *then we have the stronger result that G itself can be expressed as a lexicographically ordered direct sum of ordered subgroups of rank one,* i.e., if $0 = G_0 < G_1 < \cdots < G_\rho = G$ is the sequence of isolated subgroups of G then G contains ordered subgroups $H_1, H_2, \cdots, H_\rho$ of rank one such that $G_{\rho-i} = H_{i+1} \oplus H_{i+2} \oplus \cdots \oplus H_\rho$ where the sum is lexicographically ordered direct (i.e., every element g of $G_{\rho-i}$ has a unique expression $g = h_{i+1} + h_{i+2} + \cdots + h_\rho$ with h_j in H_j, if $g^* = h^*_{i+1} + h^*_{i+2} + \cdots + h^*_\rho$ with h^*_j in H_j is any other element of $G_{\rho-i}$ different from g and if $h_{i+j} = h^*_{i+j}$ for $j = 1, 2, \cdots, t-1$ and $h_{i+t} \neq h^*_{i+t}$ then $g < g^*$ or $g > g^*$ according as $h_{i+t} < h^*_{i+t}$ or $h_{i+t} > h^*_{i+t}$). For $\rho = 1$ this is obvious and to apply induction, assume that $\rho > 1$ and that the assertion is true for $\rho - 1$. Let H be the maximal isolated subgroup of G. Then H is of rank $\rho - 1$, G/H is of rank one, and H and G/H are both rationally complete. By the induction hypothesis, $H = H_2 \oplus H_3 \oplus \cdots \oplus H_\rho$ where the sum is lexicographically ordered direct. Let B be a rationally independent rational basis of G/H and for each b in B fix β in G contained in the residue class b. Let H_1 be the set of all elements of G which depend rationally on the β's. Then it can be easily verified that

* Received February 2, 1956.

H_1 is order isomorphic to G/H and that $G = H_1 \oplus H = H_1 \oplus H_2 \oplus \cdots \oplus H_\rho$ where the direct sums are lexicographically ordered.

In the general case when G is not necessarily rationally complete, it is enough to observe that G can always be embeded in a rational completion i. e. in an abelian group G^* which is rationally complete and for any element g^* of G we have $ng^* \varepsilon G$ for some nonzero integer n. The ordering of G can be uniquely extended to G^* and then G^* is again of rank ρ. As shown above, we can find in G^* subgroups $H_1, H_2, \cdots, H_\rho$ of rank one such that $G^* = H_1 \oplus H_2 \oplus \cdots \oplus H_\rho$ where the direct sum is lexicographically ordered. Now each H_i being archimedian is order isomorphic to a subgroup of R and hence G^* is order isomorphic to a subgroup of $R^{(\rho)}$. Therefore G is order isomorphic to a subgroup of $R^{(\rho)}$. Now we shall show that if G is not rationally complete then it need not be order isomorphic to a lexicographically ordered direct sum of rank one groups. To show this, let G denote the torsionfree abelian group studied by L. Pontrijagin in Example 2 of Appendix 1 of his paper: The theory of topological commutative groups; *Annals of Mathematics*, Vol. 35 (1934), pp. 361-388. Then G is of rational rank two and it is not expressible as a direct sum of rational rank one subgroups. Let (u, v) be a rational basis of G. Then each element g of G has a unique expression $g = au + bv$ where a and b are rational numbers. Let $g^* = a^*u + b^*v$ be any other element of G where a^* and b^* are rational numbers. We set $g > g^*$ either if $a > a^*$ or if $a = a^*$ and $b > b^*$. This turns G into an ordered abelian group and we have $\rho(G) = 2$. Suppose if possible that G is the lexicographically ordered direct sum of two ordered rank one subgroups H_1 and H_2. Then we have that $r(H_1) = r(H_2) = 1$ and hence that G is the direct (group theoretic) sum of two rational rank one subgroups; this is absurd.

We observe that if in the above proof we replace simple induction by transfinite induction then we obtain the following more general result: Let G be an ordered abelian group and let $S(G)$ be the family of isolated subgroups of G simply ordered by inclusion. Assume that $S(G)$ is well ordered. Let G^* be a rational completion of G and extend the ordering of G to G^* (uniquely). Then G^* is order isomorphic to a lexicographically ordered direct sum of a well ordered family T [which is order isomorphic to $S(G)$] of archimedian groups.

HARVARD UNIVERSITY.

REFERENCES.

[1] S. Abhyankar and O. Zariski, "Splitting of valuations in extensions of local domains," *Proceedings of the National Academy of Sciences*, vol. 41 (1955), pp. 84-90.

[2] ———, "Local uniformization on algebraic surfaces over ground fields of characteristic $p \neq 0$," forthcoming in the *Annals of Mathematics*.

[3] I. S. Cohen, "On the structure and ideal theory of complete local rings," *Transactions of the American Mathematical Society*, vol. 59 (1946), pp. 54-106.

[4] W. Krull, "Allgemeine Bemertungstheorie," *Jour. für Mathematik*, vol. 167 (1932), pp. 160, 196.

[5] ———, *Idealtheorie*, Berlin (1935).

[6] ———, "Dimensionstheorie in Stellenringen," *Jour. für Mathematik*, vol. 179 (1938), pp. 204-226.

[7] ———, "Zur Theorie der kommutativen Integritätsbereiche," *ibid.*, vol. 192 (1954), pp. 230-252.

[8] S. MacLane and O. F. G. Schilling, "Zero-dimensional branches of rank one on algebraic varieties," *Annals of Mathematics*, vol. 40 (1939), pp. 507-520.

[9] F. K. Schmidt, "Über die Erhaltung der Kettensätze der Idealtheorie bei beliebigen endlichen Körpererweiterungen," *Mathematische Zeitschrift*, vol. 41 (1936), pp. 443-440.

[10] O. Zariski, "Foundations of a general theory of birational correspondences," *Transactions of the American Mathematical Society*, vol. 53 (1943), pp. 490-542.

[11] ———, "Reduction of the singularities of algebraic three dimensional varieties," *Annals of Mathematics*, vol. 45 (1944), pp. 472-542.

[12] ———, "The concept of a simple point of an abstract algebraic variety," *Transactions of the American Mathematical Society*, vol. 62 (1947), pp. 1-52.

[13] ———, "A simple analytical proof of a fundamental property of birational transformations," *Proceedings of the National Academy of Sciences*, vol. 35 (1949), pp. 62-66.

[14] ———, "Applicazioni geometriche della teoria delle valutazioni," *Rendiconti del Circolo Matematico di Palermo*, Ser. 5, vol. 13 (1954), pp. 1-38.

ON FRENET'S EQUATIONS.*

By Aurel Wintner.

1. If Γ is an oriented, rectifiable Jordan arc in the X-space, where $X = (x, y, z)$, let $\Gamma \varepsilon C''$ mean that the vector function $X(s)$ has a continuous second derivative with respect to the arc length s. Thus $U_1 = X'$, where $' = d/ds$, is a continuously differentiable unit vector. Let κ denote the (non-negative, continuous) function $| U_1' |$ of s, the curvature on Γ.

Let Γ be called a *Frenet curve* (in symbols: $\Gamma \varepsilon F$) if it is a $\Gamma \varepsilon C''$ corresponding to which there exists a vector function U_3 possessing the following properties: U_3 is a unit vector which, at every s, is orthogonal to U_1 and has a continuous derivative U_3' which is linearly dependent on the vector product $[U_3, U_1]$.

2. The following considerations will deal with the class $\Gamma \varepsilon F$. The emphasis will be two-fold: on the one hand, $\Gamma \varepsilon F$ does not involve the existence of a third derivative for the vector function $X(s)$, where Γ: $X = X(s)$ and, on the other hand, $\Gamma \varepsilon F$ allows the vanishing of the curvature $\kappa(s) = | X''(s) |$ (on arbitrarily complicated s-sets, which must, of course, be closed sets, since $\kappa(s)$ is continuous). It is therefore unexpected that $\Gamma \varepsilon F$ proves to be the natural condition for the existence of a Frenet theory. In fact, it turns out that

(I) *every* $\Gamma \varepsilon F$ *possesses a unique, continuous torsion* $\tau = \tau(s)$
and that

(II) *Frenet's equations are valid on every* $\Gamma \varepsilon F$.

(I) and (II) imply the existence and the continuous differentiability of a principal normal and of a binormal on every $\Gamma \varepsilon F$. *But these unit vectors in Frenet's equations need not be unique* (although, by (I), the torsion *is* unique), not even if the differentiability assumption $\Gamma \varepsilon C''$, contained in the assumption $\Gamma \varepsilon F$, is refined to the existence of arbitrarily high derivatives of $X(s)$; simply because Γ can contain segments of straight lines (in a finite or infinite number).

The generality under which (I) and (II) are secured by it (Sections 3-4

* Received July 27, 1955.

below) is not, however, the only merit of the *curve* class $\Gamma \varepsilon F$. In fact, its principal merit consists in its ability to deal with certain desiderata in the theory of *surfaces*. This will be explained and carried out in Sections 5 and 6-7, respectively.

3. Let a Γ satisfy the conditions required of a $\Gamma \varepsilon F$ in Section 1. Define $U_2 = U_2(s)$ by placing $U_2 = [U_3, U_1]$. Then (U_1, U_2, U_3) is a continuously differentiable matrix function of s, and is an orthogonal matrix, of deterimnant $+1$.

By the last of the requirements which define the condition $\Gamma \varepsilon F$, the (continuous) vector U_3' is a scalar multiple of U_2. Let this scalar factor be denoted by $-\tau$, and let $\tau = \tau(s)$ be declared to be the torsion of Γ. Then, since $U_3' = -\tau U_2$ is continuous, as is the unit vector U_2, it is clear that τ is continuous. It also follows from $|U_2| = 1 \neq 0$ that τ is uniquely determined by U_3' and U_2 and, therefore, by U_2. But $U_2 = U_2(s)$ is not uniquely determined by Γ when $\kappa(s) = 0$. The following consideration will, however, show that τ is unique at every s, as claimed by (I), Section 2 (in fact, $U_3' = -\tau U_2$ implies that $\tau = -U_3 \cdot U_2'$, whereas U_2 and/or U_3 becomes indeterminate only at points contained in the *interior* of such straight line segments as may be contained in Γ; but $\tau = \tau(s)$ proves to be $\equiv 0$ on any such segment).

Let (F_i), where $i = 1, 2, 3$, denote the i-th of the relations

$$U_1' = \kappa U_2, \qquad U_2' = -\kappa U_1 + \tau U_3, \qquad U_3' = -\tau U_2.$$

Then (F_3) holds by the definition of τ. If (F) denotes the system of the three relations (F_i), then the assertion of (II), Section 2, is that (F) holds at every point s of every $\Gamma \varepsilon F$. This assertion will be proved in Section 4.

4. Suppose first only that $\Gamma \varepsilon C''$. Then there exists a continuous $\kappa = \kappa(s) \geqq 0$, defined as the length of the continuous vector function X''. Denote by $B_1, B_2, \cdots$ the (finite or infinite) sequence of open s-intervals on which $U_1'(s) \neq 0$, where $U_1 = X'$. The s-interval, say A, which is the topological map of the entire arc Γ is supposed to be closed. It can be assumed that $A - B$, where $B = B_1 + B_2 + \cdots$, is not vacuous. For otherwise $\kappa \equiv 0$ on A, hence Γ is a line segment, and so all three equations (F_i) become satisfied by placing $\tau \equiv 0$, $U_3 = [U_1, U_2]$ and choosing U_2 to be any constant unit vector which is orthogonal to $U_1 = \text{const}$. Correspondingly, it will be clear that it is sufficient to prove (II), Section 2, under the assumption that Γ contains no line segments.

Then the interval A is the closure of the open set B. But $U_2 = U_1'/\kappa$ defines a $U_2 = U_2(s)$ on every component B_i of B. Since κ and U_1' are

defined and continuous on A, and since B is dense on A, it is clear that it is possible to define on A a unique continuous U_2 satisfying (F_1), the first of the three equations (F).

On the other hand, (F_3) is true by definition (Section 3). Finally, (F_2) follows by inserting (F_1) and (F_3) into the derivative of $U_2 = [U_3, U_1]$.

5. Let $\Gamma \varepsilon F^*$ mean that Γ belongs to the subclass F^* of the class F which is restricted by the hypothesis that $\kappa(s) \neq 0$ on Γ. The books of analytically-minded authors do not fail to emphasize that the (continuous) torsion, which they define as $1/\kappa^2$ times the determinant of the first three derivatives of X, is left undefined unless $\kappa(s) \neq 0$ on Γ *and* $X(s)$ possesses a (continuous) third derivative. The resulting geometrical interpretation of the torsion (in terms of the binormal spherical image) is classical, of course. The analytical advantages of this geometrical approach were noticed, however, only recently, in [2], pp. 770-774.

It was shown in [4], pp. 243-244, that a $\Gamma \varepsilon C''$ satisfies the C^3-assumption of the traditional treatment if and only if it is a $\Gamma \varepsilon F^*$ on which $\kappa(s) > 0$, instead of being just continuous, is continuously differentiable, a hypothesis which prevents the formulation of important geometrical situations, for which $\Gamma \varepsilon F^*$ turns out to be the natural assumption (cf. [4], p. 246, pp. 247-249 and p. 257). This holds not only for a curve Γ as such but also for classical curves Γ drawn on the surface, such as asymptotic lines and geodesics (cf. [2], pp. 773-774, and [3], pp. 608-610).

It is clear, however, that not only the traditional presentation (C^3) but also its geometrical refinement (F^*) excludes every straight line Γ, the (otherwise) isolated zeros s of $\kappa(s)$ if $\Gamma \varepsilon A$, and closed s-sets (which can be Cantor sets of positive measure) if $S \varepsilon C^\infty$, where A and C^∞ denote the classes of curves Γ for which the function $X(s)$ is analytic and possesses derivatives of arbitrarily higher order, respectively. This leads to the anomalous situation that, for instance, such fundamental facts as center around the concept of geodesic torsion or the Beltrami-Enneper theorem have never been formulated for those points of Γ at which $\kappa(s) = 0$.

This was the reason for replacing the F^*-class by the F-class, as introduced in Section 1. It remains to be shown that the class $\Gamma \varepsilon F$ is not too inclusive to be useful in disposing of the anomalies referred to before.

6. Let n be a positive integer and S a (sufficiently small, open, simply connected piece of a) surface in the X-space, where $X = (x, y, z)$. Then an $S \varepsilon C^n$ is defined by the property that S has some parametrization S: $X = X(u, v)$ in which all partial derivatives of the function $X(u, v)$ which have a (collective) order not exceeding n exist, are continuous, and those of

the first order are such that their vector product $[X_u, X_v]$ does not vanish at any point of the (u,v)-domain under consideration.

The most immediate use of the concept of a continuous torsion, as defined in Section 3, is that the following assertion becomes true without any restriction:

(i) *If* Γ *is a geodesic on an* $S\,\varepsilon\, C^2$, *then* $\Gamma\,\varepsilon\, F$.

In order to prove (i), let $N = N(s)$ denote the (oriented) unit normal vector of S along Γ. Since $S\,\varepsilon\, C^2$ and since Γ is a geodesic, $\Gamma\,\varepsilon\, C''$. Hence $U_1 = X'$ has a continuous derivative. The same is true of U_2 and U_3, if the latter two vectors are defined to be N and $[U_1, U_2] = [X', N]$, respectively. But the definition of a geodesic Γ requires the vanishing of $[X', N]' \cdot X'$ along Γ, whilst $[X', N]' \cdot [X', N]$ vanishes identically (since $[X', N] \cdot [X', N] = 1$). This means that U_3' is orthogonal to U_1 as well as to U_3 and is, therefore, a scalar multiple of U_2. Hence, (i) follows from the definition of the class $\Gamma\,\varepsilon\, F$ (Section 1).

In answering a question I raised some time ago, Professor Hartman ([1], pp. 724-726) has shown that if D is a direction through P within T, where P denotes a point of any $S\,\varepsilon\, C^2$ and T the plane tangent to S at P, then (P, D) determines a *unique* geodesic of S. Hence, (i) supplies the following result:

(ii) *If* $S\,\varepsilon\, C^2$, *then every point of* S *and a direction through it* (*on* S) *define a unique geodesic torsion,* the latter being defined as the torsion (at P), supplied by (i), of the geodesic determined by (P, D).

If $\Gamma\,\varepsilon\, C'$ means that Γ is an oriented, rectifiable Jordan arc for which $X(s)$ possesses a continuous first derivative, then a corollary of (ii) can be formulated as follows:

(iii) *A* $\Gamma\,\varepsilon\, C'$ *on an* $S\,\varepsilon\, C^2$ *is a line of curvature of* S *if and only if the geodesic torsion of* Γ *vanishes identically* (*provided that* S *is free of umbilical points*).

In fact, since (ii) assures the existence of a geodesic torsion along Γ, it is only necessary to repeat the considerations of [3], pp. 608-609, in order to obtain (iii).

7. Let $S\,\varepsilon\, C^3$. Then the Gaussian curvature K and the coefficients of the second fundamental form $Ldu^2 + 2Mdudv + Ndv^2$ (when referred to a C^3-parametrization $X = X(u,v)$ of S) are continuously differentiable functions of (u,v). Suppose that K is negative on S. Then $LN - M^2 < 0$. Hence the identical vanishing of the second fundamental form along a curve

Γ of S represents for Γ either of two systems of ordinary differential equations and, if either of them is written in the form $du/dt = f(u,v)$, $dv/dt = g(u,v)$, then both functions f, g are continuously differentiable and do not contain the independent variable, t. Hence the solutions $u = u(t)$, $v = v(t)$ have continuous second derivatives with respect to t. Since an asymptotic line Γ of S is defined by $\Gamma: X = X(u(t), v(t))$, it follows that $\Gamma \varepsilon C''$. It will be shown that $\Gamma \varepsilon C''$ can be improved to $\Gamma \varepsilon F$:

(iv) *If $K < 0$ on an $S \varepsilon C^3$, then there exists a continuous torsion on every asymptotic line Γ of S; in fact, $\Gamma \varepsilon F$. In addition, the square of the torsion is identical with $-K$ on S.*

First, since $S \varepsilon C^3$ implies that $S \varepsilon C^2$, it follows from the fact $\Gamma \varepsilon C''$ (which could *not* have been concluded from just $S \varepsilon C^2$) that the unit normal $N = N(s)$ to S has a continuous first derivative along Γ, as does $U_1 = X'(s)$, where Γ: $X = X(s)$. Hence, if U_2 and U_3 are defined to be $[N, X']$ and N, respectively, then (U_1, U_2, U_3) is a continuously differentiable orthogonal matrix (of determinant $+1$). But the definition of an asymptotic line Γ, used above, requires the vanishing of $X' \cdot N'$ along Γ, whilst $N \cdot N'$ vanishes identically (since $N \cdot N = 1$). This means that $U_3' = N'$ is orthogonal to U_1 as well as to U_3 and is, therefore, a scalar multiple of U_2. Hence the assertion $\Gamma \varepsilon F$ of (iv) follows from the definition of the Frenet class F (Section 1).

The remaining assertion of (iv), that concerning the value of $-K$ along Γ (Beltrami-Enneper), can be concluded from the Frenet system (F), belonging to $U_3 = N$, in the usual way. Cf. the corresponding remarks in [2], p. 773, which deal with (iv) under the (by now superfluous) hypothesis that the curvature $\kappa(s)$ does not vanish on the asymptotic line Γ. Actually, the last assertion of (iv) then follows from the first assertion of (iv) without the assumption $\kappa(s) > 0$ also, simply for reasons of continuity (in fact, if the trivial case $\kappa(s) \equiv 0$ is disregarded, then the zeros $s = s_0$ of $\kappa(s)$ are cluster points of an open s-set on which $\kappa(s) > 0$). Cf. also (III) in [3], p. 609.

It may finally be mentioned that, in view of [5], pp. 858-859, the assumption, $S \varepsilon C^3$, in (iv) appears to be necessary and sufficient in order that an $S \varepsilon C^2$ of negative K be such as to possess a C^2-parametrization S: $X = X(u,v)$ in which the parameter lines $u =$ const., $v =$ Const. are asymptotic lines of S.

8. If $\Gamma \varepsilon C'''$ means that Γ: $X = X(s)$ is a rectifiable Jordan arc for which the vector function $X(s)$ possesses a continuous third derivative, then the following criterion (*) holds:

(*) *A* $\Gamma \varepsilon F$ *is a* $\Gamma \varepsilon C'''$ *if and only if its curvature* $\kappa(s)$ *is continuously differentiable.*

This is a generalization of a criterion in [4], pp. 243-244, where (*) was proved under the hypothesis $\kappa(s) \neq 0$ (that is, under the assumption that $\Gamma \varepsilon F$ is strengthened to $\Gamma \varepsilon F^*$ in the sense of Section 5). It is, however, clear from the proof (*loc. cit.*) and from the definition of the Frenet class F (Section 1) that the first of the assertions of (*), the assertion in which the continuous differentiability of $\kappa(s)$ is the assumption, holds also if $\Gamma \varepsilon F^*$ is relaxed to $\Gamma \varepsilon F$.

In order to prove the second assertion of (*), suppose that $\Gamma \varepsilon C'''$ and let A and $B = B_1 + B_2 + \cdots$ be defined as in Section 4. Then $U_1 = X'$ has a continuous second, hence $\kappa^2 = X''.X''$ a continuous first, derivative on A. Actually, $\kappa = \kappa(s) \geqq 0$ itself must have a continuous first derivative on A. In order to prove this at an arbitrary point s_0 of A, two cases must be distinguished, according as s_0 is in B or in $A - B$. In the first case, the assertion is trivial, since $\kappa(s_0) > 0$. Let therefore $\kappa(s_0) = 0$. Then, if (F_1) is applied at s_0 and at a nearby s, it is seen that $\kappa(s)U_2(s)$ is identical with the difference $U_1'(s) - U_1'(s_0)$. But the ratio of the latter to $s - s_0$ tends, as $s \to s_0$, to a limit, since $U_1(s)$ has a second derivative. Consequently, $\kappa(s)U_2(s)/(s - s_0)$ must tend, as $s \to s_0$, to a limit. Since $U_2(s) \to U_2(s_0) \neq 0$, this means that $\kappa(s)/(s - s_0)$ has a limit. In view of $\kappa(s_0) = 0$, this proves the differentiability of $\kappa(s)$ at s_0, and therefore at every point of A. Finally, the continuity of the derivative $\kappa'(s)$ follows from the circumstance that, since $U_1''(s)$ is continuous, the preceding limit process holds uniformly.

It is clear from Section 2 that (*) can be interpreted as a criterion supplying, in terms of the behavior of the curvature, a necessary and sufficient condition in order that a $\Gamma \varepsilon C''$ be a $\Gamma \varepsilon C'''$, provided that Γ has a continuous torsion. In fact, this proviso is contained in (and, when combined with $\Gamma \varepsilon C''$, becomes equivalent to) the hypothesis $\Gamma \varepsilon F$ of (*).

Appendix.

With reference to a Γ: $X(s)$ of class F, let R denote either sheet ($0 < t < \infty$ or $-\infty < t < 0$) of the ruled surface generated by the binormal of F; so that R: $X = X(t, s) \equiv tU_3(s)$, where $t \neq 0$ and $X = (x, y, z)$. Thus $X(t, s)$ is a function of class C^1 and

$$[X_t, X_s] = [U_3, tU_3'] = -t\tau[U_3, U_2],$$

by (F_3). Hence the unit normal vector, say $M = M(t, s)$, of R exists if $\tau \neq 0$. In fact, $\pm M$ is seen to be $[U_2, U_3] = U_1$. Accordingly, if $\Gamma \varepsilon F$ and $\tau \neq 0$, then R is a surface of class C^1. It turns out, however, that *R is a surface*

of class C^2 (even though its "geometrical" parametrization, R: $X = tU_3(s)$, cannot in general be of class C^2, since $U_3(s) \varepsilon C^2$ is surely not true if only $\Gamma \varepsilon F$, or for that matter $\Gamma \varepsilon C^3$, is assumed). This can be seen as follows:

The surface $R \varepsilon C^1$ has a C^1-parametrization in terms of which the unit normal M is a function of class C^1. Hence, $R \varepsilon C^2$ can be concluded by an elementary argument (based on the classical theorem on local implicit functions) which was repeatedly applied in the writings of Hartman and myself (cf., e.g., [*3*], pp. 368-369, where, incidentally, the issue involved, that of the preservation of the C^2-character of a surface $S \varepsilon C^2$ under a "parallel" deformation, is very similar to the present construction of the surface R from the curve Γ, where $\Gamma \varepsilon F$, hence $\Gamma \varepsilon C''$). Of course, $R \varepsilon C^2$ means that *some* parametrization $X = X(u;v)$ of R is of class C^2 (with $[X_u, X_v] \neq 0$). The point is that, against expectation, the "geometrical" parametrization, that in which $(u,v) = (t,s)$, fails to be such a parametrization in general.

Needless to say, the ruled surface R is a torse, in the sense of having a normal $M = M(t,s)$ which is independent $(= \pm U_1)$ of the position t on any generating line ($s =$ const.) of R. Since $R \varepsilon C^2$, this implies * that the Gaussian curvature K of R vanishes identically; cf. the end of the footnote below. The result is therefore as follows:

(*a*) *If* $\tau \neq 0$ *on a* $\Gamma \varepsilon F$ (*where* $\kappa \neq 0$ *is not assumed*), *then the ruled surface* $R = R(\Gamma)$: $X = tU_3(s)$, *where* $0 < |t| < \infty$, *is a torse* $(K \equiv 0)$ *of class* C^2.

---

* Certain difficulties inherent to Euler's definition of a torse are known since Lebesgue's thesis ([*2*], pp. 319-342). But it may not be necessary to go to such extremes as Lebesgue went (continuous but not one-to-one parametrizations) in order to show that the theory of torses is not as simple as it appears from the texts of differential geometry, including the rigor-conscious books. For is it true that if $K \equiv 0$ on an $S \varepsilon C^2$, then a neighborhood of every point of S can be "ruled," so as to be a torse in Euler's sense also? I can neither prove nor believe this, not even under the assumption $S \varepsilon C^\infty$ which, in view of the possibility of clustering zeros of H (i.e., of "flat" points, where $H^2 = 0 = K$), is hardly stronger than $S \varepsilon C^2$ (in view of Theorem (†), p. 134, of [*1*], there is no trouble when a torse $S \varepsilon C^2$ is free of flat points, but there could be trouble even if there is just one such point). A counterexample, with $S \varepsilon C^\infty$, would be the first such instance in the differential geometry of surfaces as to require the full force of (function-theoretical) analyticity (or at least quasi-analyticity, rather than just C^∞-character).

The converse inference, that in which the identical vanishing of the Gaussian curvature is the assertion, can be concluded without any calculation (whenever the surface is of class C^2). In fact, if S is a torse in Euler's synthetic sense of the term, then the normal image of each of the generating lines is a single point. Since the C^2-character of S suffices to justify the applicability of Fubini's theorem on product measures, it thus becomes clear that the normal image of any subset of S is of measure zero on the unit sphere.

This result, (α), has a counterpart, (β), for the case in which the binormal U_3 is replaced by the tangent U_1 of Γ (but not for the case of the remaining U_i, the principal normal; the hindrance being that equation (F_2), in contrast to (F_3) and (F_1), does not contain just one U on the right). The counterpart is as follows:

(β) *If* $\kappa \neq 0$ *on a* $\Gamma \varepsilon F$ (*where* $\tau \neq 0$ *is not assumed*), *i.e., if* $\Gamma \varepsilon F^*$, *then the ruled surface* $P = P(\Gamma)$: $tU_1(s)$, *where* $0 < |t| < \infty$, *is a torse* ($K \equiv 0$) *of class* C^2.

In fact, what corresponds to the last formula line when R is replaced by P is

$$[X_t, X_s] = [U_1, tU_1'] = t\kappa[U_1, U_2],$$

by (F_1). Hence it is clear that (β) follows by a repetition of the proof of (α).

It will be noted that the proof of (α) or (β) succeeds because the ruled surfaces traditionally attached to a Γ: $X = X(s)$, which are the ruled surfaces $X(s) + tU_i(s)$, are reduced to $tU_i(s)$, i.e., that the base curve, instead of being Γ, is made to be 0.

THE JOHNS HOPKINS UNIVERSITY.

REFERENCES.

[1] P. Hartman, "On the local uniqueness of geodesics," *American Journal of Mathematics*, vol. 72 (1950), pp. 723-730.

[2] ——— and A. Wintner, "On the fundamental equations of differential geometry," *ibid.*, vol. 72 (1950), pp. 757-774.

[3] ——— and A. Wintner, "On geodesic torsions and parabolic and asymptotic curves," *ibid.*, vol. 74 (1952), pp. 607-625.

[4] A. Wintner, "On the infinitesimal geometry of curves," *ibid.*, vol. 75 (1953), pp. 241-259.

[5] ———, "On indefinite binary Riemannian metrics," *ibid.*, vol. 77 (1955), pp. 853-867.

APPENDIX.

[*1*] P. Hartman and A. Wintner, "On the curvatures of a surface," *American Journal of Mathematics*, vol. 75 (1953), pp. 127-141.

[*2*] H. Lebesgue, "Integrale, longueur, aire," *Annali di Matematica Pura ed Applicata*, ser. 3, vol. 7 (1902), pp. 231-359.

[*3*] A. Wintner, "On parallel surfaces," *American Journal of Mathematics*, vol. 74 (1952), pp. 365-376.

THE STRUCTURE OF FACTORS OF AUTOMORPHY.*

By R. C. Gunning.

One of the principal tools in the study of automorphic functions of one and several complex variables is the representation of an automorphic function as the quotient of two holomorphic functions whose zeros are invariant under the group of transformations acting, although the functions themselves are not invariant; these auxiliary functions, that is, satisfy a relation of the form $f(Tz) = \nu_T(z) f(z)$, where $\nu_T(z)$ are holomorphic and nowhere vanishing functions called factors of automorphy. Such auxiliary functions were first studied in connection with elliptic functions in one variable and abelian functions in several variables, where they have been called the Jacobi or intermediary functions. Their importance in deriving the Riemann conditions on the period matrix of a multi-torus, in discussing the divisors on a multi-torus, and in proving the existence theorems of abelian functions, has long been recognized. Appell was the first to prove, in several variables, that the particular factors of automorphy involved in the Jacobi functions are, in a sense, the most general possible factors on the multi-torus [1]. Corresponding auxiliary functions have been used in other cases also, in connection with the Poincaré and Eisenstein series, but, as Bochner has pointed out [3], no systematic attempts have been made to give a general classification of these factors corresponding to the work of Appell. The present paper contains the proofs of some results announced earlier [11] in connection with this classification problem, and a discussion of the significance of this classification as a generalization of that of the Jacobi functions.

The method utilized in the proofs is the application of potential theory in the form of the theory of harmonic differential forms on Kaehler manifolds; recent investigations of this subject, in connection with analytic manifolds and transcendental algebraic geometry, have yielded a wealth of applicable results. Although some knowledge of the terminology and notation used in the study of complex manifolds is presupposed, the relevant topological and differential-geometric results are collected in the first two sections. It has been possible to develop the subject on this foundation without requiring any

* Received December 3, 1955.

extensive algebraic or topological background; the presentation given here uses this approach for the benefit of those whose interest lies in the function-theoretic aspects of the subject. A supplementary section, Section 5, illustrates an algebraic development of the classification theorem.

I would like to express my warmest thanks here to Professor Bochner for suggesting this problem to me, and for many valuable discussions on this and related topics.

I. Introductory.

1. Suppose that D is a simply-connected complex analytic manifold of complex dimension n, and that Γ is a countable group of analytic homeomorphisms of D onto itself subject to the following restrictions:

(1.1) Each point $z \in D$ lies in a coordinate neighborhood U such that either $TU = U$ or $TU \cap U = \emptyset$ for any $T \in \Gamma$.

(1.2) For each pair of points $z_1, z_2 \in D$ either $\Gamma z_1 = \Gamma z_2$ or there exist coordinate neighborhoods U_1, U_2 of z_1, z_2 respectively such that $\Gamma U_1 \cap \Gamma U_2 = \emptyset$. (Here, $\Gamma z = \bigcup_{T \in \Gamma} Tz$.)

(1.3) The quotient space D/Γ, with the quotient topology, is a compact space.

(1.4) There is a Kaehler metric on D which is invariant under the action of the group Γ.

For any subset $W \subset D$ let Γ_W be the subgroup of Γ consisting of all transformations T for which $TW = W$; thus whenever W is a coordinate neighborhood, Γ_W is a properly discontinuous group of analytic homeomorphisms on a bounded affine subdomain. Applying the standard methods of the theory of automorphic functions on bounded affine domains to the particular coordinate neighborhoods (1.1), one sees that the subgroups Γ_z are finite for each z, and that there are arbitrarily small coordinate neighborhoods of each point which satisfy (1.1); in particular there is a coordinate neighborhood U of each point $z \in D$ such that for any $T \in \Gamma$, either $TU \cap U = \emptyset$ or $TU = U$ and $Tz = z$.

The collection of all points of D left fixed by a transformation $T \in \Gamma$ other than the identity forms a proper analytic subvariety S_T of D. The set

$$S = \bigcup_{T \in \Gamma - I} S_T,$$

which one notes is a Γ-invariant analytic subvariety of D, is called the *singular set* of D; its image S/Γ under the canonical projection $\rho: D \to D/\Gamma$

is called the *singular set* of D/Γ. The mapping $\rho: D-S \to D/\Gamma - S/\Gamma$ is then a local homeomorphism, defining the structure of a complex analytic manifold on $D/\Gamma - S/\Gamma$ and exhibiting $D-S$ as a regular covering manifold. Select a fixed base point $z_0 \epsilon D-S$, and for each $T \epsilon \Gamma$ select a path α_T from z_0 to Tz_0 in $D-S$; the uniquely defined homotopy class of the loop $\rho(\alpha_T)$ in $D/\Gamma - S/\Gamma$ will be denoted by $\tilde{\psi}(T)$, so that $\tilde{\psi}$ is the standard isomorphism of Γ onto the fundamental group $\pi_1(D/\Gamma - S/\Gamma)$ of $D/\Gamma - S\Gamma$ based at z_0. The injection of $D/\Gamma - S/\Gamma$ into D/Γ induces a further homomorphism of $\pi_1(D/\Gamma - S/\Gamma)$ onto $\pi_1(D/\Gamma)$, and the composite of this homomorphism with $\tilde{\psi}$ defines a homomorphism ψ of Γ onto $\pi_1(D/\Gamma)$. One can see that the kernel of this homomorphism is the normal subgroup Γ_0 of Γ generated by all transformations which possess fixed points, so that $\Gamma/\Gamma_0 \cong \pi_1(D/\Gamma)$; furthermore, Γ_0 is contained in the normal subgroup of Γ generated by all transformations which are of finite order. Any homomorphism of Γ into a commutative field therefore vanishes on Γ_0 and on all commutators of elements of Γ, so determines a homomorphism on $H_1(D/\Gamma)$; for the field of real numbers in particular, the so-called universal coefficient theorem [7] asserts that each such homomorphism corresponds to a one-dimensional cohomology class on D/Γ with real coefficients.

Suppose that the group Γ is presented in some fixed manner as the factor group of a finitely generated free group H modulo a group of relations P, which is also finitely generated. The free generators of H will be denoted by $t_1, \cdots, t_p$, and their images under the canoncial projection homomorphism $\mathrm{H} \to \Gamma$ will be denoted by $T_1, \cdots, T_p$ respectively. The subgroup $[\mathrm{P}, \mathrm{H}]$ of H generated by all words of the form $s^{-1}r^{-1}sr$ for arbitrary $r \epsilon \mathrm{P}$, $s \epsilon \mathrm{H}$, is called the *commutator* of P in H; replacing P by H itself, one defines correspondingly the *commutator subgroup* $[\mathrm{H}, \mathrm{H}]$ of H. Note that $[\mathrm{P}, \mathrm{H}]$ is a normal subgroup of H, hence also of the group $\mathrm{P} \cap [\mathrm{H}, \mathrm{H}]$ defined as the set-theoretic intersection of P and $[\mathrm{H}, \mathrm{H}]$. Letting $H_2(D)$, $H_2(D/\Gamma)$ denote the singular homology groups of D and D/Γ, the projection $\rho: D \to D/\Gamma$ induces a homomorphism $\rho_*: H_2(D) \to H_2(D/\Gamma)$ whose image will be denoted by $S_2(D/\Gamma)$. We shall next construct explicitly a useful homomorphism

$$\Phi: \mathrm{P} \cap [\mathrm{H}, \mathrm{H}]/[\mathrm{P}, \mathrm{H}] \to H_2(D/\Gamma)/S_2(D/\Gamma), \tag{2}$$

related to a theorem of H. Hopf [13].

An element $v \epsilon \mathrm{P} \cap [\mathrm{H}, \mathrm{H}]$ is uniquely expressible as a reduced word in the free generators of H, say

$$v = t_{\alpha(q)}{}^{\epsilon(q)} \cdots t_{\alpha(1)}{}^{\epsilon(1)},$$

where $\epsilon(\nu) = \pm 1$. Let $j(1)$ be the smallest integer for which

$$t_{\alpha(j(1))}^{\epsilon(j(1))} = t_{\alpha(1)}^{-\epsilon(1)};$$

let $j(2)$ be the smallest integer distinct from 1 and $j(1)$ such that

$$t_{\alpha(j(2))}^{\epsilon(j(2))} = t_{\alpha(2)}^{-\epsilon(2)}; \text{ etc.}$$

Thus there is associated to the word v a unique one-to-one mapping j of a subset of the integers $1, \cdots, q$ onto the complementary set. For each r in the domain of the mapping j construct an arbitrary singular 1-simplex σ_r^1 in $D - S$ such that

$$\partial\sigma_r^1 = T_{\alpha(j(r))}^{\epsilon(j(r))} \cdots T_{\alpha(1)}^{\epsilon(1)} z_0 - T_{\alpha(r-1)}^{\epsilon(r-1)} \cdots T_{\alpha(1)}^{\epsilon(1)} z_0.$$

Introduce also the singular 1-simplex $\bar{\sigma}_r^1 = -T_{\alpha(r)}^{\epsilon(r)}\sigma_r^1$ and the singular 1-chain $\kappa^1(v) = \sum_r (\sigma_r^1 + \bar{\sigma}_r^1)$, where the summation is extended over all r in the domain of the mapping j. It is clear from the construction that $\partial\kappa^1(v) = 0$ and that $\rho\kappa^1(v) = 0$, where $\rho : D \rightarrow D/\Gamma$ is again the canonical projection. Since D is simply-connected, there is a singular 2-chain $\kappa^2(v)$ such that $\partial\kappa^2(v) = \kappa^1(v)$, and hence $\partial\rho\kappa^2(v) = \rho\partial\kappa^2(v) = 0$. Let $\phi(v)$ be the coset of the homology class of $\rho\kappa^2(v)$ in $H_2(D/\Gamma)/S_2(D/\Gamma)$. Two types of arbitrary choices were involved in the definition of the elements $\phi(v)$, namely the selections of the σ_r^1 and of the $\kappa^2(v)$; clearly the result is independent of these choices, so that ϕ is a well-defined mapping. One sees immediately that it is even a homomorphism of $\mathrm{P} \cap [\mathrm{H}, \mathrm{H}]$ into $H_2(D/\Gamma)/S_2(D/\Gamma)$.

One can further see that the homomorphism ϕ is actually independent of the choice of the auxiliary mapping j, so long as j is one-to-one between a subset of the indexing set and its complement, and $t_{\alpha(r)}^{\epsilon(r)} = t_{\alpha(j(r))}^{-\epsilon(j(r))}$. As a first corollary of this independence, one notes that $[\mathrm{P}, \mathrm{H}]$ lies in the kernel of the homomorphism ϕ; the induced homomorphism Φ on the factor group $\mathrm{P} \cap [\mathrm{H}, \mathrm{H}]/[\mathrm{P}, \mathrm{H}]$ is the desired homomorphism.

In case Γ has no fixed points, D/Γ is itself a complex analytic manifold and D is its universal covering manifold. The mapping Φ is then an isomorphism onto; indeed, as a second corollary of the independence of Φ from the auxiliary mapping j, one sees that Φ coincides with the isomorphism established by Hopf. The group $S_2(D/\Gamma)$ may also be interpreted in this case as the image of the second homotopy group of D/Γ in the second homology group under the canonical mapping.

2. Some differential-geometric properties of the spaces D and D/Γ are also needed as background for the subsequent discussion. These are all

known theorems for manifolds without singularities; the fact that they remain true for manifolds with singularities of the types admitted here in the spaces D/Γ when Γ contains transformations with fixed points has been verified in detail by W. Baily [2]. For the notation to be used in the sequel, as well as for definitions of the fundamental concepts involved, see for example [10, 18].

By *de Rham's theorems*, the cohomology groups of the space D/Γ with complex coefficients are isomorphic to the d-cohomology groups of Γ-invariant complex differential forms on the manifold D, where d is the exterior differentiation operator on D. *Hodge's theorem* asserts the existence of a unique Δ-harmonic Γ-invariant differential form representing each cohomology class. Since the manifold D is complex-analytic, the ordinary exterior differentiation may be decomposed as the sum of two complex operators, $d = \partial + \bar{\partial}$. One may define a harmonic operator $\square$, associated to the differentiation $\bar{\partial}$, analogous to the Laplace-Beltrami operator Δ; since D is Kaehler, one even has $\square = \frac{1}{2}\Delta$. Analogs of the Hodge theorem may be obtained for the operator $\square$. A crucial step in this development is the *decomposition theorem*: every Γ-invariant C^∞ differential form ϕ of type (p, q) on D which satisfies $\bar{\partial}\phi = 0$ may be written in the form $\phi = \bar{\partial}\psi + \theta$, where ψ is a Γ-invariant differential form of type $(p, q-1)$ and θ is a Γ-invariant $\square$-harmonic form of type (p, q), in the sense that $\square\theta = 0$. From the Kaehler hypothesis, one secures $d\theta = 0$. From this follows a topological invariance theorem: if b^r is the r-th Betti number of the space D/Γ and $h^{p,q}$ is the dimension of the complex linear space of Γ-invariant $\square$-harmonic differential forms of type (p, q), then $b^r = \sum_{p+q=r} h^{p,q}$.

The Γ-invariant $\square$-harmonic differential forms $\omega_\alpha(z)$ of type $(1, 0)$ on D are called the *abelian differentials* (of the first kind) on D/Γ; these may be defined equivalently as the Γ-invariant differential forms of type $(1, 0)$ which satisfy either $\bar{\partial}\omega_\alpha(z) = 0$ or $d\omega_\alpha(z) = 0$. Since the forms $\omega_\alpha(z)$ are closed, it is possible to introduce in addition the *abelian integrals* $w_\alpha(z) = \int^z \omega_\alpha(\zeta)$; these may be characterized as the set of all single-valued holomorphic functions on D satisfying $w_\alpha(Tz) = w_\alpha(z) + \hat{\omega}_\alpha(T)$ for all $T \in \Gamma$, where $\hat{\omega}_\alpha(T)$ are complex constants called the *periods* of the integral. If $\{\omega_\alpha(z)\}$ form a basis for the complex linear space of abelian differentials on D/Γ, then since the operator $\square$ is real on a Kaehler manifold, the conjugate differentials $\{\bar{\omega}_\alpha(z)\}$ form a basis for the complex linear space of $\square$-harmonic differential forms of type $(0, 1)$; applying the topological invariance theorem, in the case $r = 1$, there are precisely $\frac{1}{2}b^1$ linearly independent abelian differentials on D/Γ. Let $\{T_j\}$ be a set of transformations of Γ representing the infinite cyclic generators of the abelianized group $\Gamma/[\Gamma, \Gamma]$; from our previous geo-

metrical considerations, there are at most b^1 such elements. These may be called a *basis* for the group Γ. Setting $\omega_{\alpha j} = \hat{\omega}_\alpha(T_j)$, the matrix $\Omega = (\omega_{\alpha j})$ is called the *period matrix* corresponding to the bases $\{\omega_\alpha(z)\}$ and $\{T_j\}$. Since the mapping $\hat{\omega}_\alpha : T \rightarrow \hat{\omega}_\alpha(T)$ is a homomorphism of Γ into the additive group of complex numbers, it vanishes on elements of finite order and on the subgroup $[\Gamma, \Gamma]$; hence all the periods $\hat{\omega}_\alpha(T)$ of the abelian differential $\omega_\alpha(z)$ are integer linear combinations of the basic periods listed in row α of the period matrix. Since there can exist no abelian differential with purely imaginary periods, one notes that there are precisely b^1 basic transformations, and that the b^1 by b^1 *matrix* $\begin{pmatrix} \Omega \\ \bar{\Omega} \end{pmatrix}$ is non-singular.

II. Classification of Factors of Automorphy.

3. The collection of all holomorphic functions on the complex analytic manifold D form an abelian group $\mathcal{A}(D)$ under addition. The group Γ then has a natural interpretation as a group of operators on $\mathcal{A}(D)$, where the action of an element $T \in \Gamma$ on $f(z) \in \mathcal{A}(D)$ is defined to yield $f(Tz)$.

Definition. A *summand of automorphy* for the group Γ on D is a mapping σ of Γ into $\mathcal{A}(D)$, the image of an element $T \in \Gamma$ being denoted by $\sigma_T(z)$, such that $\sigma_{ST}(z) = \sigma_S(Tz) + \sigma_T(z)$. In algebraic terminology, an equivalent restatement of this definition is that a summand of automorphy is a one-cocycle of Γ with coefficient group $\mathcal{A}(D)$ [8].

LEMMA 1. *There exists for any summand of automorphy σ a C^∞ function $f(z)$ on D such that*

$$f(Tz) = f(z) + \sigma_T(z). \tag{3}$$

Proof. Select a finite number of pairs of open coordinate neighborhoods $V_j \subset \bar{V}_j \subset U_j$ of D such that the sets $\rho(V_j)$ cover D/Γ, that for any $T \in \Gamma$ either $TU_j = U_j$ or $TU_j \cap U_j = \phi$, and that the subgroups $\Gamma_j = \Gamma_{U_j}$ are of orders $m_j < \infty$. For each j construct a real C^∞ function $\tilde{\mu}_j(z)$ on D such that $0 \leqq \tilde{\mu}_j(z) \leqq 1$, that $\tilde{\mu}_j(z) = 0$ for $z \notin U_j$, and that $\tilde{\mu}_j(z) = 1$ for $z \in V_j$. Then the functions

$$\mu_j(z) = \sum_{T \in \Gamma} \tilde{\mu}_j(Tz) / \sum_{T \in \Gamma} \sum_j \tilde{\mu}_j(Tz)$$

are Γ-invariant C^∞ functions on D, $\mu_j(z) = 0$ whenever $z \notin \bigcup_{T \in \Gamma} TU_j$, and $\sum_j \mu_j(z) \equiv 1$.

The functions $\sigma_{jT}(z) = \mu_j(z)\sigma_T(z)$ form a set of C^∞, although not holomorphic, summands of automorphy. Let

$$f_j(z) = \begin{cases} -1/m_j \sum_{S \in \Gamma_j} \sigma_{jS}(z) & \text{for } z \in U_j, \\ f_j(T^{-1}z) + \sigma_{jT}(T^{-1}z) & \text{for } z \in TU_j, \\ 0 & \text{otherwise.} \end{cases}$$

Then $f_j(z)$ is clearly C^∞ on D and satisfies $f_j(Tz) = f_j(z) + \sigma_{jT}(z)$. The function $f(z) = \sum_j f_j(z)$ is then the function whose existence was to be demonstrated.

In general there will not exist a holomorphic function satisfying (3). Select any C^∞ function $f(z)$ such that $f(Tz) = f(z) + \sigma_T(z)$. Since $\sigma_T(z)$ are holomorphic, $\bar{\partial}\sigma_T(z) = 0$ and $\bar{\partial}f(Tz) = \bar{\partial}f(z)$; thus $\bar{\partial}f(z)$ is a Γ-invariant $\bar{\partial}$-closed C^∞ differential form of type $(0,1)$. Applying the decomposition theorem of Section 2, $\partial f(z) = \bar{\partial}f^*(z) + \theta_\sigma(z)$, where $f^*(z)$ is a Γ-invariant C^∞ function on D, and $\theta_\sigma(z)$ is a Γ-invariant harmonic differential form of type $(0,1)$ on D. The form $\theta_\sigma(z)$, or the one-dimensional cohomology class $\hat{\theta}_\sigma$ it represents, is an obstruction associated to the summand $\sigma_T(z)$. It depends only upon the summand of automorphy $\sigma_T(z)$. For if $f_1(z)$ is another C^∞ function satisfying (3), and $\bar{\partial}f_1(z) = \bar{\partial}f_1^*(z) + \theta_\sigma^1(z)$, set $f_0(z) = f(z) - f^*(z) - f_1(z) + f_1^*(z)$ and $\theta_\sigma^0(z) = \theta_\sigma(z) - \theta_\sigma^1(z)$; then $f_0(Tz) = f_0(z)$ and $\bar{\partial}f_0(z) = \theta_\sigma^0(z)$ is harmonic, which implies $\theta_\sigma^0(z) = 0$. Clearly there will exist a holomorphic function satisfying (3) if and only if the obstruction $\theta_\sigma(z)$ of the summand $\sigma_T(z)$ vanishes; this in turn can be achieved by a simple modification of the original summand.

THEOREM 1. *If σ is any summand of automorphy for the group Γ on D, there is a unique homomorphism $\hat{a}: T \to a_T$ of Γ into the additive group of real numbers for which there will exist a holomorphic function $g(z)$ satisfying $g(Tz) = g(z) + \sigma_T(z) + 2\pi i\, a_T$.*

Proof. Replacing the function $f(z)$ considered in the above paragraph by $f(z) - f^*(z)$ if necessary, one may assume that $f(Tz) = f(z) + \sigma_T(z)$ and that $\theta_\sigma(z) = \bar{\partial}f(z)$ is a harmonic differential form. Thus

$$0 = d\theta_\sigma(z) = \partial\bar{\partial}f(z) = -\bar{\partial}\partial f(z) = -d\partial f(z),$$

so that $\partial f(z)$ is a closed, holomorphic differential form of type $(1,0)$. Since D is simply-connected, $g(z) = \int^z \partial f(\zeta)$ is a well-defined holomorphic function on D for which $dg(z) = \partial f(z)$. Consequently

$$g(Tz) = g(z) + \sigma_T(z) + 2\pi i\, b_T \tag{4}$$

for some complex constants b_T.

The mapping $T \to b_T$ is clearly a homomorphism of Γ into the additive group of complex numbers, so that $b_T = 0$ whenever T lies in the commutator subgroup of Γ, or is an element of finite order. As a result, this homomorphism may be considered as defining a one-dimensional cohomology class $\hat{b}_\sigma$ of D/Γ with complex coefficients. One already has the one-dimensional cohomology class $\hat{\theta}_\sigma$ defined by the secondary obstruction of the summand σ. For any 1-cycle α of the space D/Γ, represented by a singular 1-chain of D with boundary $Tz_0 - z_0$, one secures

$$\begin{aligned}\hat{\theta}_\sigma(\alpha) &= \int_{z_0}^{Tz_0} \theta_\sigma(z) = \int_{z_0}^{Tz_0} \bar{\partial} f(z)\\ &= \int_{z_0}^{Tz_0} df(z) - \partial f(z)\\ &= f(Tz_0) - g(Tz_0) - f(z_0) + g(z_0)\\ &= -2\pi i\, b_{T_0} = -2\pi i\, \hat{b}_\sigma(\alpha);\end{aligned}$$

therefore $\hat{b}_\sigma = \frac{1}{2\pi i}\hat{\theta}_\sigma$.

The most general holomorphic function satisfying an equation of the desired type (4), with perhaps different complex constant terms a_T appearing, is given by

$$g^*(z) = g(z) + \sum_{\alpha=1}^{\frac{1}{2}b^1} \xi_\alpha w_\alpha(z) + a$$

where a and ξ_α are complex constants and $\{w_\alpha(z)\}$ form a basis for the abelian integrals on D/Γ. Letting $\{T_j\}$ be a basis for the group Γ in the sense of Section 2 and $\Omega = (\omega_{\alpha j})$ be the corresponding period matrix, all of the constants a_T are expressible as linear combinations, with integer coefficients, of the numbers

$$a_{T_j} = b_{T_j} + \sum_{\alpha=1}^{\frac{1}{2}b^1} \xi_\alpha \omega_{\alpha j}.$$

Since the matrix $\begin{pmatrix}\Omega\\ \bar{\Omega}\end{pmatrix}$ is non-singular, values ξ_α may be selected so that a_{T_j}, and hence all a_T, are real numbers. Then the function $g^*(z)$ is obviously unique up to the additive constant a, while the homomorphism $T \to a_T$ is completely unique. This homomorphism defines a one-dimensional cohomology class $\hat{a}_\sigma$ of D/Γ with real coefficients. Letting $\hat{\omega}_\alpha$ be the one-dimensional cohomology classes of D/Γ defined by the abelian differentials $\omega_\alpha(z)$,

$$\hat{a}_\sigma = -\frac{1}{2\pi i}\hat{\theta}_\sigma + \sum_{\alpha=1}^{\frac{1}{2}b^1} \xi_\alpha \hat{\omega}_\alpha.$$

The remarks made during the course of the preceding proof have given a significant further interpretation to the homomorphism $T \rightarrow a_T$ by considering the one-dimensional cohomology class $\hat{a}_\sigma$ defined by this homomorphism, by way of the isomorphism $\Gamma/\Gamma_0 \cong \pi_1(D/\Gamma)$ established in Section 1. The cohomoloy classes $\hat{\omega}_\alpha$ form a subgroup of $H^1(D/\Gamma, C)$ which may be called the subgroup of *analytic cohomology classes*. The class $\hat{a}_\sigma$ may then be described as the unique real cohomology class in that coset of the subgroup of analytic cohomology classes which contains the obstruction $-\frac{1}{2\pi i}\hat{\theta}_\sigma$ of the summand σ. Also of interest is the cohomology class $\chi_2(\sigma) = \exp 2\pi i\, \hat{a}_\sigma$ with coefficients in the one-dimensional unitary group.

4. The collection of all holomorphic, nowhere-vanishing functions on the complex manifold D form an abelian group $\mathfrak{M}(D)$ under multiplication. The group Γ then has a natural interpretation as a group of operators on $\mathfrak{M}(D)$, where the action of an element $T \in \Gamma$ on $f(z) \in \mathfrak{M}(D)$ is defined to yield $f(Tz)$.

Definition. A *factor of automorphy* for the group Γ on D is a mapping ν of Γ into $\mathfrak{M}(D)$, the image of an element $T \in \Gamma$ being denoted by $\nu_T(z)$, such that $\nu_{ST}(z) = \nu_S(Tz)\nu_T(z)$. The set of all factors of automorphy form an abelian group under multiplication.

In algebraic terminology, a factor of automorphy is a one-cocycle of the group Γ with coefficients in the group $\mathfrak{M}(D)$. Two such cocycles or factors μ and ν are *cohomologous* if there exists a holomorphic, nowhere-vanishing function $h(z)$ on D such that $h(Tz)/h(z) = \mu_T(z)/\nu_T(z)$. We shall develop a classification of these cohomology classes analogous to that given in Theorem 1 for the corresponding additive cocycles, or summands of automorphy. It is more convenient, and actually more natural, to approach this problem in two steps, the first of which consists in the study of a weaker classification of factors of automorphy. For this purpose, one recalls that a *character* of the group is a homomorphism $\hat{c}: T \rightarrow c_T$ of Γ into the multiplicative group of complex numbers of modulus 1, the one-dimensional unitary group. Two factors of automorphy μ and ν are *equivalent* if there exists a character $\hat{c} = \{c_T\}$ of the group Γ and a holomorphic non-vanishing function $h(z)$ on D such that $\mu_T(z) = c_T\nu_T(z)\, h(Tz)/h(z)$.

For any group H and any abelian group G, let $\operatorname{Hom}(\mathrm{H}; G)$ be the group of all homomorphisms of H into G; for any subgroups $\mathrm{H}_1 \subset \mathrm{H}$ and $G_1 \subset G$, let $\operatorname{Hom}(\mathrm{H}_1, \mathrm{H}; G_1, G)$ be the subgroup of $\operatorname{Hom}(\mathrm{H}_1; G_1)$ consisting of all elements which can be extended to homomorphisms of H into G. As in

Section 1, the group Γ is assumed presented in some manner as the quotient of a finitely generated free group H modulo a group of relations P; the free generators of H are $t_1, \cdots, t_p$, and their images in Γ are $T_1, \cdots, T_p$ respectively.

If ν is any factor of automorphy, select a branch of the logarithm $\sigma_{t_j}(z) = \log \nu_{T_j}(z)$ for each free generator t_j. After this selection has been made, a unique holomorphic function $\sigma_t(z)$ can be associated to each $t \in \mathrm{H}$ by requiring that $\sigma_{st}(z) = \sigma_s(Tz) + \sigma_t(z)$. For every $t \in \mathrm{H}$, $\nu_T(z) = \exp \sigma_t(z)$; consequently whenever $t \in \mathrm{P}$

$$\hat{\sigma}(t) = \tfrac{1}{2\pi i}\sigma_t(z) \tag{5}$$

is an integer. Furthermore for $s, t \in \mathrm{P}$, $\hat{\sigma}(st) = \hat{\sigma}(s) - \hat{\sigma}(t)$, while for $t \in \mathrm{P}$ and $v \in \mathrm{H}$, $\hat{\sigma}(vtv^{-1}) = \frac{1}{2\pi i}[\sigma_v(TV^{-1}z) + \sigma_t(V^{-1}z) - \sigma_v(V^{-1}z)] = \hat{\sigma}(t)$; this therefore defines an element $\hat{\sigma} \in \mathrm{Hom}(\mathrm{P}/[\mathrm{P},\mathrm{H}]; Z)$, where Z is the additive group of integers. If $\sigma^*_{t_j} = \log \nu_{T_j}(z)$ are defined by selecting different branches of the logarithms, then for any $t \in \mathrm{H}$, $\sigma^*_t(z) - \sigma_t(z) = 2\pi i\, m_t$ for some integer m_t; the mapping $\hat{m}: t \to m_t$ is a homomorphism of H into the additive group Z. That is, for any two homomorphisms $\hat{\sigma}$ and $\hat{\sigma}^*$ determined by the same factor of automorphy,

$$\hat{\sigma}^* - \hat{\sigma} \in \mathrm{Hom}(\mathrm{P}/[\mathrm{P},\mathrm{H}], \mathrm{H}/[\mathrm{P},\mathrm{H}]; Z, Z). \tag{6}$$

LEMMA 2. *If* H *is a finitely generated free group and* P *is a normal subgroup, then there is a canonical isomorphism into*

$$\phi: \frac{\mathrm{Hom}(\mathrm{P}/[\mathrm{P},\mathrm{H}]; Z)}{\mathrm{Hom}(\mathrm{P}/[\mathrm{P},\mathrm{H}], \mathrm{H}/[\mathrm{P},\mathrm{H}]; Z, \bar{Q})} \to \mathrm{Hom}(\mathrm{P} \cap [\mathrm{H},\mathrm{H}]/[\mathrm{P},\mathrm{H}]; Z),$$

where Z *is the additive group of integers and* $\bar{Q}$ *is a group containing the additive group of rational numbers.*

Proof. An element $\sigma \in \mathrm{Hom}(\mathrm{P}/[\mathrm{P},\mathrm{H}]; Z)$ may be considered as a homomorphism of P into Z such that $\sigma([\mathrm{P},\mathrm{H}]) = 0$. Let $\phi(\sigma)$ be the restriction of σ to the subgroup $\mathrm{P} \cap [\mathrm{H},\mathrm{H}] \subset \mathrm{P}$; this defines a canonical homomorphism ϕ of $\mathrm{Hom}(\mathrm{P}/[\mathrm{P},\mathrm{H}]; Z)$ into $\mathrm{Hom}(\mathrm{P} \cap [\mathrm{H},\mathrm{H}]/[\mathrm{P},\mathrm{H}]; Z)$. Obviously $\mathrm{Hom}(\mathrm{P}/[\mathrm{P},\mathrm{H}], \mathrm{H}/[\mathrm{P},\mathrm{H}]; Z, \bar{Q})$ lies in the kernel of ϕ. Conversely suppose that $\phi(\sigma) = 0$, or what is the same, that $\sigma(\mathrm{P} \cap [\mathrm{H},\mathrm{H}]) = 0$; to complete the proof it is only necessary to show that σ can be extended to a homomorphism of H into Q. Let P be generated by $u_1, u_2, \cdots$, and write

$$u_j = t_1^{m_{j1}} \cdots t_p^{m_{jp}} s_j$$

for some $s_j \in [\mathrm{H}, \mathrm{H}]$ and integers m_{jk}. Values $\sigma(u_j)$ are given, and values $x_k = \sigma(t_k)$ are to be selected in such a manner that σ is a homomorphism on H; this simply amounts to solving the system of linear equations

$$\sigma(u_j) = m_{j1}x_1 + \cdots + m_{jp}x_p$$

for some rational numbers x_k. If there is a linear dependence among the right-hand members of these equations, say $\sum_j n_j m_{jk} = 0$ for $1 \leqq k \leqq p$, and some integers n_j, then

$$\prod_j u_j^{n_j} \in \mathrm{P} \cap [\mathrm{H}, \mathrm{H}],$$

and consequently $\sum_j n_j \sigma(u_j) = \sigma(\prod_j u_j^{n_j}) = 0$. These equations are thereby consistent, and do admit solutions.

Definition. The *character class* (associated to the presentation $\Gamma \cong \mathrm{H}/\mathrm{P}$) of a factor of automorphy ν is the element

$$\chi_1(\nu) = \phi(\hat{\sigma}) \in \mathrm{Hom}(\mathrm{P} \cap [\mathrm{H}, \mathrm{H}]/[\mathrm{P}, \mathrm{H}]; Z),$$

where $\hat{\sigma}$ is any homomorphism of the form (5) and ϕ is the canonical isomorphism of Lemma 2. As a consequence of (6), the character class is uniquely determined by the factor of automorphy ν alone. If μ and ν are two factors of automorphy, then $\chi_1(\mu\nu) = \chi_1(\mu) + \chi_1(\nu)$; thus χ_1 is a homomorphism of the group of factors of automorphy into $\mathrm{Hom}(\mathrm{P} \cap [\mathrm{H}, \mathrm{H}]/[\mathrm{P}, \mathrm{H}]; Z)$. The extent to which the character class is independent of the presentation $\Gamma \cong \mathrm{H}/\mathrm{P}$ will be discussed in the supervening section, in which the underlying algebraic structure will be examined.

LEMMA 3. *Two sets of factors of automorphy μ and ν have the same character class if and only if there exists a character $\hat{c} = \{c_T\}$ of the group Γ and a summand of automorphy σ such that $\mu_T(z) = c_T \nu_T(z) \exp \sigma_T(z)$.*

Proof. Introducing the quotient factor $\eta = \mu/\nu$, the factors μ and ν have the same character class if and only if $\chi_1(\eta) = 0$. Corresponding to the presentation $\Gamma \cong \mathrm{H}/\mathrm{P}$, select holomorphic functions $\tau_t(z)$ for each $t \in \mathrm{H}$ such that $\tau_{st}(z) = \tau_s(Tz) + \tau_t(z)$ and that $\eta_T(z) = \exp \tau_t(z)$; this may be done in the manner in which we defined the character class, for example. Any character $\hat{c} = \{c_T\}$ of the group Γ can be written in the form $c_T = \exp 2\pi i\, a_t$, where $a_t = \hat{a}(t)$ and

$$\hat{a} \in \mathrm{Hom}(\mathrm{P}/[\mathrm{P}, \mathrm{H}], \mathrm{H}/[\mathrm{P}, \mathrm{H}]; Z, \bar{Q}) \tag{7}$$

for $\tilde{Q}$ the additive group of real numbers. Letting $\sigma_t(z) = \tau_t(z) - 2\pi i\, a_t$, it follows that

$$\mu_T(z)/\nu_T(z) = \eta_T(z) = c_T \exp \sigma_t(z).$$

The functions $\sigma_t(z)$ will define a summand of automorphy for a suitable choice of $\hat{a}$ if and cnly if $\sigma_t(z) \equiv 0$ for all $t \in \mathrm{P}$, or what is the same, if and only if $\hat{\sigma} = 0$ for the homomorphism $\hat{\sigma}$ defined by (5). Since $\hat{\sigma} = \hat{\tau} - \hat{a}$, it is possible to select $\hat{a}$ such that $\hat{\sigma} = 0$ if and only if $\hat{\tau}$ belongs to the group (7) containing all possible $\hat{a}$; by Lemma 2, this in turn is equivalent to the fact that $\phi(\hat{\tau}) = \chi_1(\eta) = 0$.

THEOREM 2. *Two factors of automorphy are equivalent if and only if they have the same character class.*

Proof. By Lemma 3, factors μ and ν have the same character class if and only if there is a character $\hat{c} = \{c_T\}$ of Γ and a summand of automorphy σ such that $\mu_T(z) = c_T \nu_T(z) \exp \sigma_T(z)$. By Theorem 1, there exists for any summand σ a holomorphic function $g(z)$ such that $g(Tz) = g(z) + \sigma_T(z) + 2\pi i\, a_T$ for some real numbers a_T. Letting $b_T = c_T \exp[-2\pi i\, a_T]$ and $h(z) = \exp g(z)$, μ and ν have the same character class if and only if $\mu_T(z) = b_T \nu_T(z) h(Tz)/h(z)$, which was to be proved.

The final cohomological classification of factors of automorphy follows trivially. From each class of equivalent factors of automorphy select a basic element ν; any other factor in the same equivalence class can be written in the form $\mu_T(z) = c_T \nu_T(z) \exp \sigma_T(z)$ for some summand σ. The mapping $T \to c_T$ is actually a homomorphism on the abelianized group $\Gamma/[\Gamma,\Gamma]$; one then sees trivially that c_T and σ may be modified in such a manner that $c_T = 1$ except on the torsion elements of $\Gamma/[\Gamma,\Gamma]$. The homomorphism $T \to c_T$ is then unique, and will be called the torsion class $T_\nu(\mu)$ of μ with respect to ν; the class $\chi_2(\sigma) = \chi_{2,\nu}(\mu)$ of the summand σ may also be considered as an invariant associated to the factor μ with respect to ν. The three invariants $\chi_1(\mu)$, $T_\nu(\mu)$, and $\chi_{2,\nu}(\mu)$ then completely characterize the cohomology classes of factors.

5. The preceding analysis clearly demonstrated that the equivalence classification of factors of automorphy is an algebraic consequence of Theorem 1. As a short digression, we shall rephrase the above discussion in a more abstract setting which emphasizes these formal aspects of the argument. Let Γ be a group, which for the sake of convenience we shall continue to assume

finitely generated, and G be its group ring $Z(\Gamma)$. Consider the following collection of G-modules and G-homomorphisms:

$$\begin{tikzcd}
0 \arrow[r] & J \arrow[r, "j"] & A \arrow[r, "\phi"] & M \arrow[r] & 0 \\
0 \arrow[r] & J \arrow[r, "k"] \arrow[u, "i"] & P \arrow[r, "\psi"] \arrow[u, "\alpha"] & R \arrow[r] \arrow[u, "\beta"] & 0
\end{tikzcd}$$

The two horizontal rows are exact sequences, the diagram is commutative, and i is the identity map. From this one derives the diagram:

$$\begin{tikzcd}
\cdots \arrow[r] & H^1(\Gamma, J) \arrow[r, "j^*"] & H^1(\Gamma, A) \arrow[r, "\phi^*"] & H^1(\Gamma, M) \arrow[r, "\delta_1^*"] & H^2(\Gamma, J) \arrow[r] & \cdots \\
\cdots \arrow[r] & H^1(\Gamma, J) \arrow[r, "k^*"] \arrow[u, "i^*"] & H^1(\Gamma, P) \arrow[r, "\psi^*"] \arrow[u, "\alpha^*"] & H^1(\Gamma, \mathrm{R}) \arrow[r, "\delta_2^*"] \arrow[u, "\beta^*"] & H^2(\Gamma, J) \arrow[r] \arrow[u, "i^*"] & \cdots
\end{tikzcd}$$

The two horizontal rows are again exact sequences, the diagram is commutative, and i^* is the identity map. It follows immediately from this diagram that whenever α^* is an isomorphism onto, then β^* is an isomorphism which has as image the subgroup $\delta_1{}^{*-1}i^*\delta_2{}^*[H^1(\Gamma, J)] \subset H^1(\Gamma, M)$. That is, if we define a homomorphism $\chi : H^1(\Gamma, M) \to H^2(\Gamma, J)/i^*\delta_2{}^*[H^1(\Gamma, R)]$ by associating to each $m \in H^1(\Gamma, M)$ the coset containing $\delta_1{}^*(m)$, then the image of β^* is the kernel of χ. Two 1-cocycles μ and ν of the group Γ with coefficients in M are therefore cohomologous to the image under β^* of a 1-cocycle c having coefficients in R if and only if $\chi(\mu) = \chi(\nu)$; the 1-cocycle c is unique when it exists.

Returning to our particular case, set $J = Z$, $A = \mathcal{A}(D)$, $M = \mathcal{M}(D)$, $P =$ additive group of real numbers, $R =$ one-dimensional unitary group, j and $k =$ injection maps, and ϕ and $\psi =$ maps: $x \to \exp 2\pi i x$. Theorem 1 states that α^* is an isomorphism onto; hence two factors μ and ν are equivalent if and only if $\chi(\mu) = \chi(\nu)$. Applying Lemma 2, one sees that

$$H^2(\Gamma, Z)/i^*\delta_2{}^*[H^1(\Gamma, R)] \cong \mathrm{Hom}(\mathrm{P} \cap [\mathrm{H}, \mathrm{H}]/[\mathrm{P}, \mathrm{H}]; Z)$$

whenever Γ is presented as the factor group of a finitely generated free group H modulo a group of relations P, and thus identifies χ under this isomorphism with the character class.

III. The Role of the Character Class.

6. A *relatively automorphic function* associated to a factor of automorphy ν is a meromorphic function $f(z)$ on D such that $f(Tz)=\nu_T(z)f(z)$. The set of all relatively automorphic functions associated to the factor ν form a complex linear space $\mathcal{L}(\nu)$. If μ is a factor of automorphy cohomologous to ν, there exists a function $h(z)$ holomorphic and nowhere-vanishing on D such that $h(Tz)=h(z)\mu_T(z)/\nu_T(z)$; then whenever $f(z)\in\mathcal{L}(\nu)$, the function $T_{\mu\nu}f(z)=h(z)f(z)\in\mathcal{L}(\mu)$. The mapping $T_{\mu\nu}:\mathcal{L}(\nu)\to\mathcal{L}(\mu)$ so defined is an *analytic isomorphism*; that is, in addition to being an isomorphism of the complex linear spaces, it satisfies the following two conditions: (i) $T_{\mu\nu}$ preserves the structure of the zeros and poles of the functions involved, in the sense that $T_{\mu\nu}f(z)/f(z)$ is holomorphic and nowhere-vanishing, (ii) $T_{\mu\nu}$ preserves the representation of automorphic functions, in the sense that an automorphic function represented as $f(z)=f_1(z)/f_2(z)$ for $f_j(z)\in\mathcal{L}(\nu)$ is also represented as $f(z)=T_{\mu\nu}f_1(z)/T_{\mu\nu}f_2(z)$. Conversely it is clear that whenever $\mathcal{L}(\mu)$ and $\mathcal{L}(\nu)$ are non-vacuous and analytically isomorphic, the factors μ and ν are cohomologous. This expresses the function-theoretic significance of the concept of cohomologous factors of automorphy.

A singular 1-simplex σ^1 of D is in *general position* with respect to a function $f(z)$ meromorphic on D if $f(z)$ is finite-valued and non-zero at each point of the support of σ^1; a singular 1-chain κ^1 of D is in general position with respect to $f(z)$ if each component simplex in κ^1 is in general position. The set of all singular 2-chains κ^2 of D whose boundaries are in general position with respect to $f(z)$ form an abelian group, and $f(z)$ defines a linear function on this group by

$$c(f)[\kappa^2]=\frac{1}{2\pi i}\int_{\partial\kappa^2}\frac{df(z)}{f(z)}.$$

THEOREM 3. *Let χ_1 be the character class associated to a presentation $\Gamma\cong \mathrm{H}/\mathrm{P}$ and Φ be the homomorphism* (2). *Then for any factor of automorphy ν and associated relatively automorphic function $f(z)$,*

$$c(f)[\Phi(v)]=\chi_1(\nu)[v] \tag{8}$$

for every $v\in\mathrm{P}\cap[\mathrm{H},\mathrm{H}]$.

Proof. Reverting to the notation of Section 1, it is clear that to each $v\in\mathrm{P}\cap[\mathrm{H},\mathrm{H}]$ we may associate one of the standard singular 2-chains $\kappa^2(v)$

whose boundary $\kappa^1(v) = \partial\kappa^2(v)$ is in general position with respect to $f(z)$; the precise meaning of (8) is then

$$c(f)[\kappa^2(v)] = \chi_1(\nu)[v]$$

whenever $\kappa^2(v)$ is in general position with respect to $f(z)$. It follows from the definition of $\kappa^1(v)$ that

$$\int_{\kappa^1(v)} df(z)/f(z) = \sum_r \int_{\sigma_r^1 - \bar{\sigma}_r^1} df(z)/f(z)$$
$$= -\sum_r \int_{\sigma_r^1} d\nu[T_{\alpha(r)}{}^{\epsilon(r)}](z)/\nu[T_{\alpha(r)}{}^{\epsilon(r)}](z),$$

where we have written $\nu[T](z)$ for $\nu_T(z)$. Introducing functions

$$\sigma_r(z) = \log \nu[T_{\alpha(r)}{}^{\epsilon(r)}](z)$$

one secures

$$\int_{\kappa^1(v)} df(z)/f(z) = -\sum_r [\sigma_r(T_{\alpha(j(r))}{}^{\epsilon(j(r))}\cdots T_{\alpha(1)}{}^{\epsilon(1)}z_0)$$
$$-\sigma_r(T_{\alpha(r-1)}{}^{\epsilon(r-1)}\cdots T_{\alpha(1)}{}^{\epsilon(1)}z_0)]$$
$$= \sigma_v(z_0) = 2\pi i\chi_1(\nu)[v].$$

Thus

$$c(f)[\kappa^2(v)] = \frac{1}{2\pi i}\int_{\kappa^1(v)} df(z)/f(z) = \chi_1(\nu)[v], \tag{9}$$

which was to be proved.

For an interpretation of this theorem in slightly different terms, let us recall that a *divisor* $\mathfrak{D}$ on the manifold D is a finite formal sum $\mathfrak{D} = \sum_j m_j V_j$, where V_j are irreducible $(n-1)$-dimensional complex analytic subvarieties of D and m_j are positive or negative integers. In the open sets U_λ of a sufficiently fine covering of D, each subvariety V_j may be represented as the locus of the zeros of a function $d_{j\lambda}(z)$ which has no multiple factors at any point of U_λ; the function $d_\lambda(z) = \prod_j d_{j\lambda}(z)^{m_j}$ is called a minimal local equation of the divisor $\mathfrak{D}$ in U_λ. If $f(z)$ is a meromorphic function on D, there is a unique divisor $\mathfrak{D}(f)$ having $f(z)$ as a minimal local equation at each point; $\mathfrak{D}(f)$ is called the divisor of the function $f(z)$. A divisor $\mathfrak{D}$ with minimal local equations $d_\lambda(z)$ is Γ-invariant if $d_\lambda(Tz)/d_\mu(z)$ is holomorphic and non-vanishing in $U_\lambda \cap T^{-1}U_\mu$, for every $T \in \Gamma$. Clearly $\mathfrak{D}(f)$ is Γ-invariant whenever $f(z)$ is a relatively automorphic function; conversely, on those manifolds D with the property that every divisor is the divisor of a

global meromorphic function, every Γ-invariant divisor is the divisor of a relatively automorphic function. The manifolds D of this type include the manifolds of the particular examples considered in Chapter IV. This characterization of relatively automorphic functions also illustrates their role in the analytic representation of automorphic functions; for such manifolds D, every automorphic function can be represented as the quotient of two holomorphic relatively automorphic functions.

A point on an irreducible analytic subvariety V is called a simple point if there are local coordinates $z_1, \cdots z_n$ centered at the point, in terms of which V is locally a plane section $z_{p+1} = \cdots = z_n = 0$. The points which are not simple lie in a closed subvariety of V which is of still lower dimension, and may be called the singular set of V. Thus outside of its singular set, V may be given the structure of a complex p-dimensional manifold by introducing local coordinates $z_1, \cdots, z_p$ in the appropriate coordinate neighborhoods; the analytic structure induces a definite orientation on V in the usual manner. A (differentiable) singular 2-simplex σ^2 of D is in general position with respect to a divisor $\mathcal{D} = \sum_j m_j V_j$ if the support of σ^2 meets the point set $\bigcup_j V_j$ in finitely many interor points of σ^2, each of which is a simple point lying on but one of the subvarieties V_j, and at each of which the simplex σ^2 and set V_j meet transversally in the sense that their tangent spaces generate a full $2n$-dimensional space. A singular 2-chain κ^2 is in general position with respect to $\mathcal{D}$ if each of its simplices is. The set of all singular 2-chains in general position with respect to $\mathcal{D}$ form an abelian group, and $\mathcal{D}$ defines a linear function on this group by introducing the sum of the intersection multiplicities K. I. $(\mathcal{D}, \kappa^2)$, just as in the definition of the Kronecker index. It is clear from the Cauchy residue formula in one variable that whenever $\mathcal{D}$ is the divisor of a meromorphic function $f(z)$ on D, K. I. $(\mathcal{D}(f), \kappa^2) = c(f)[\kappa^2]$. Hence by Theorem 3 it follows that for any $f(z) \in \mathcal{L}(\nu)$,

$$\text{K. I. } (\mathcal{D}(f), \Phi(v)) = \chi_1(\nu)[v].$$

Now suppose that Γ contains no transformations with fixed points. For each singular 2-chain $\kappa^2{}_\Gamma$ on D/Γ select a 2-chain κ^2 on D such that $\rho(\kappa^2) = \kappa^2{}_\Gamma$; the fact that this is possible follows from the covering homotopy theorem for example. Then $\mathcal{D}(f)$ defines a linear function on the group of all 2-chains $\kappa^2{}_\Gamma$ of D/Γ which are in general position with respect to $\rho(\mathcal{D}(f))$ by K. I.$_\Gamma(\mathcal{D}(f), \kappa^2{}_\Gamma) =$ K. I. $(\mathcal{D}(f), \kappa^2)$, since this is clearly independent of the choice of κ^2. Let us call this function the 2-cocycle dual to

the divisor $\mathfrak{D}(f)$. The map Φ of Section 1 is an isomorphism onto, and hence induces an onto isomorphism

$$\Phi^*: \operatorname{Hom}(H_2(D/\Gamma)/S_2(D/\Gamma); Z) \to \operatorname{Hom}(\mathrm{P} \cap [\mathrm{H}, \mathrm{H}]/[\mathrm{P}, \mathrm{H}]; Z).$$

Since $H^2(D/\Gamma; Z) \cong \operatorname{Hom}(H_2(D/\Gamma); Z)$, the character class of any factor of automorphy represents a 2-dimensional integral cohomology class of D/Γ, which is aspherical in the sense that it vanishes on every spherical cycle. Theorem 3 asserts that the 2-cocycle dual to the divisor of a relatively automorphic function associated to a factor of automorphy ν represents the two-dimensional cohomology class defined by the character class of ν. Further, the divisor of any relatively automorphic function is aspherical in the above sense.

7. We shall continue to assume in this section that the group Γ contains no transformations with fixed points. Thus D/Γ is itself a complex manifold, and for the divisor $\mathfrak{D}(f)$ of any relatively automorphic function $f(z) \in \mathfrak{L}(\nu)$, $\rho(\mathfrak{D}(f))$ is a divisor on the manifold D/Γ. In order to discuss the homological properties of divisors further, we shall assume that the divisors $\rho(\mathfrak{D}(f)) = \sum_j m_j\rho(V_j)$, where the subvarieties V_j are given their natural orientations, can be expressed as singular cycles of D/Γ of class C^r $(r \geqq 1)$. This is a considerably weakened form of the strong covering theorem for analytic manifolds, which asserts that the manifold D/Γ can be covered by a simplicial complex of class C^r for arbitrary r in such a manner that $\rho(\mathfrak{D}(f))$ is a subcomplex. For further discussion of this theorem, see for example [15, 16].

It should be pointed out here that the 2-cocycle dual to the divisor $\mathfrak{D}(f)$, as defined previously, is just the cocycle dual to the singular cycle $\rho(\mathfrak{D}(f))$ in the usual singular sense. Utilizing the singular form of the de Rham representation [12], for any $(2n-2)$-dimensional cohomology class ϕ on D/Γ, represented by a Γ-invariant differential form $\phi(z)$ on D, $\phi[\mathfrak{D}] = \sum_j m_j \int_{V_j} \phi(z)$, the integration being extended over the C^1 singular cycles $\rho(V_j)$ on D/Γ. Further, letting $\chi_1(\nu)$ also denote the 2-dimensional cohomology class defined by the character class of the factor of automorphy ν and dual to $\mathfrak{D}(f)$, $\phi[\mathfrak{D}] = (\phi \cup \chi_1(\nu))[D/\Gamma]$, where $[D/\Gamma]$ is the fundamental cycle of the manifold D/Γ.

LEMMA 4. *Let W be an analytic subvariety of D/Γ of complex dimension*

$n-1$, and $\theta(z)$ be a Γ-invariant differential form of type $(n-1,0)$ on D with non-trivial support. Then

$$i^{(1-n^2)}\int_W \theta(z)\wedge\bar\theta(z)>0.$$

Proof. Consider firstly a singular simplex σ^{2n-2} of W which is disjoint from the singular set of W and contains an open set in the support of $\theta(z)$. Then complex coordinates $z_1,\cdots,z_n$, or alternatively the real coordinates $x_1,\cdots,x_{2n}$ with $z_j=x_{2j-1}+ix_{2j}$, can be introduced in an open neighborhood of σ^{2n-2} in D/Γ in such a manner that W is locally the subvariety $z_n=0$. Writing the differential form $\theta_n(z)$ as

$$\theta(z)=\sum_{j=1}^{n}\theta_{1\cdots j-1\, j+1\cdots n}(z)\,dz^1\wedge\cdots\wedge dz^{j-1}\wedge dz^{j+1}\wedge\cdots\wedge dz^n,$$

we have

$$\begin{aligned}\int_{\sigma^2}&\theta(z)\wedge\bar\theta(z)\\ &=\int_{\sigma^2}\theta_{1\cdots n-1}(z)\,\bar\theta_{1\cdots n-1}(z)\,dz^1\wedge\cdots\wedge dz^{n-1}\wedge d\bar z^1\wedge\cdots\wedge d\bar z^{n-1}\\ &=2^{n-1}i^{(n-1)(n+1)}\int_{\sigma^2}|\,\theta_{1\cdots n-1}(x)\,|^2\,dx^1\wedge\cdots\wedge dx^{2n}.\end{aligned}$$

Thus

$$i^{(1-n^2)}\int_{\sigma^2}\theta(z)\wedge\bar\theta(z)>\epsilon>0.$$

Select a barycentric subdivision of the chain W which is sufficiently fine that the integral of $i^{(1-n^2)}\theta(z)\wedge\theta(z)$ over all simplices of W which meet the singular set of W is less than ϵ in absolute value. The assertion of the lemma follows immediately.

We shall now derive from this lemma a property of the character class of factors of automorphy which admit holomorphic relatively automorphic functions, generalizing a well-known theorem of Frobenius on complex tori [9]. A factor of automorphy ν is called *positive* if it admits an associated relatively automorphic function $f(z)$ which is holomorphic on D. Let $\{\theta_\alpha\}$ be a basis for the subgroup of all complex-valued cohomology classes of dimension $n-1$ on D/Γ which can be represented in the sense of de Rham by differential forms $\{\theta_\alpha(z)\}$ of type $(n-1,0)$, and $\{\bar\theta_\alpha\}$ be the complex conjugate classes represented by the differential forms $\{\bar\theta_\alpha(z)\}$.

THEOREM 4. *Letting ν be a positive factor of automorphy with character*

class $\chi_1(\nu)$ *considered as a 2-cocycle of* D/Γ, *and* $[D/\Gamma]$ *be the fundamental cycle of the manifold* D/Γ, *the matrix*

$$i^{(1-n^2)}((\chi_1(\nu)\cup\theta_\alpha\cup\bar\theta_\beta)[D/\Gamma])_{\alpha\bar\beta}$$

is positive Hermitian. (Γ is assumed to contain no transformations with fixed points.)

Proof. Let $f(z)\in\mathcal{L}(\nu)$ be a holomorphic relatively automorphic function associated to the factor ν, and $\mathfrak{D}(f)=\sum_j m_jV_j$ be its divisor; then since $f(z)$ is holomorphic, $m_j>0$ for all j. For any arbitrary complex constants $\{\xi_\alpha\}$, not all of which are zero, $\theta(z)=\sum_\alpha \xi_\alpha\theta_\alpha(z)$ will be a Γ-invariant differential form of type $(n-1,0)$ on D which represents the cohomology class $\sum_\alpha \xi_\alpha\theta_\alpha$. Now on the one hand by Lemma 4,

$$\sum_j i^{(1-n^2)}m_j\int_{V_j}\theta(z)\wedge\bar\theta(z)>0.$$

On the other hand

$$\begin{aligned}\sum_j i^{(1-n^2)}m_j\int_{V_j}&\theta(z)\wedge\bar\theta(z)\\ &=\sum_j\sum_{\alpha,\beta} i^{(1-n^2)}m_j\xi_\alpha\bar\xi_\beta\int_{V_j}\theta_\alpha(z)\wedge\bar\theta_\beta(z)\\ &=\sum_{\alpha,\beta} i^{(1-n^2)}\xi_\alpha\bar\xi_\beta(\theta_\alpha\cup\bar\theta_\beta)[\mathfrak{D}(f)]\\ &=\sum_{\alpha,\beta} i^{(1-n^2)}\xi_\alpha\bar\xi_\beta(\chi_1(\nu)\cup\theta_\alpha\cup\bar\theta_\beta)[D/\Gamma].\end{aligned}$$

This demonstrates the theorem.

8. The method of Section 6 yields a criterion for determining which elements of $\mathrm{Hom}(\mathrm{P}\cap[\mathrm{H},\mathrm{H}]/[\mathrm{P},\mathrm{H}];Z)$ arise as possible character classes of factors of automorphy. If $g(z)$ is any C^∞ function on the manifold D for which $\bar\partial\partial g(z)$ is a Γ-invariant differential form, associate to $g(z)$ the element $\chi_g\in\mathrm{Hom}(\mathrm{P}\cap[\mathrm{H},\mathrm{H}]/[\mathrm{P},\mathrm{H}];C^+)$ defined by

$$\chi_g(v)=\int_{\kappa^2(v)}\bar\partial\partial g(z).$$

Here C^+ is the additive group of complex numbers and $\kappa^2(v)$ is one of the standard 2-cycles representing the image $\Phi(v)$ of the homomorphism Φ of Section 1; the element $\chi_g(v)$ is clearly well-defined.

THEOREM 5. *If $g(z)$ is a C^∞ function on D such that $\bar{\partial}\partial g(z)$ is Γ-invariant and χ_g is integral-valued, then some multiple of χ_g is the character class of a factor of automorphy; conversely the character class of any factor of automorphy can be so represented.*

Proof. To demonstrate the converse first, note that in the proof of Theorem 3 the second equality in (9) remains true when $df(z)/f(z)$ is replaced by any closed differential form $\rho(z)$ of degree 1 which is well-defined in a neighborhood of $\kappa^1(v)$ and satisfies $\phi(Tz) = \phi(z) + d\log \nu_T(z)$ in that neighborhood. Applying Lemma 1, select a C^∞ function $g(z)$ which is real-valued and satisfies $g(Tz) = g(z) + \frac{1}{2\pi i}\log|\nu_T(z)|^2$, where the real values of the logarithm are selected; then $\partial g(Tz) = \partial g(z) + \frac{1}{2\pi i} d\log \nu_T(z)$. Then by (9) and Stokes' theorem,

$$\chi_1(\nu)[v] = \int_{\kappa^1(v)} \partial g(z) = \int_{\kappa^2(v)} \bar{\partial}\partial g(z) = \chi_g(v)$$

for every $v \epsilon \mathrm{P} \cap [\mathrm{H}, \mathrm{H}]$.

If $\partial\partial g(z)$ is a Γ-invariant differential form on D, then for any transformation $T \epsilon \Gamma$ consider the form $\theta_T(z) = \partial g(Tz) - \partial g(z)$. Since $\partial\bar{\theta}_T(z) = \bar{\partial}\partial g(Tz) - \bar{\partial}\partial g(z) = 0$, $\theta_T(z)$ are closed, holomorphic differential forms, and it follows from their definition that $\theta_{ST}(z) = \theta_S(Tz) + \theta_S(z)$. For each free generator t_j of H select an indefinite integral $\sigma_{t_j}(z) = \int^z \theta_{T_j}(\zeta)$, and extend the functions so constructed to a collection indexed by all $t \epsilon \mathrm{H}$ and satisfying $\sigma_{st}(z) = \sigma_s(Tz) + \sigma_t(z)$. Whenever $r \epsilon \mathrm{P}$, $\hat{\sigma}(r) = \sigma_r(z)$ is a constant, and for all $s \epsilon \mathrm{H}$, $\hat{\sigma}(srs^{-1}) = \hat{\sigma}(r)$. It is again clear from (9) that $\hat{\sigma}(v) = \chi_g(v)$ for all $v \epsilon \mathrm{P} \cap [\mathrm{H}, \mathrm{H}]$. Thus if $\hat{\sigma}(r) \epsilon Z$ for all $r \epsilon \mathrm{P}$, then $\nu_T(z) = \exp \sigma_t(z)$ defines a factor of automorphy with character class $\chi_1(\nu) = \chi_g$.

Even when $\hat{\sigma}(r)$ are not always integers, it is still true by hypothesis that $\hat{\sigma}(v) = \chi_g(v) \epsilon Z$ for all $v \epsilon \mathrm{P} \cap [\mathrm{H}, \mathrm{H}]$. Moreover, applying Lemma 2 after replacing H by P and P by $\mathrm{P} \cap [\mathrm{H}, \mathrm{H}]$, there is a homomorphism α of P into the additive group of rational numbers such that $\hat{\alpha}(v) = \hat{\sigma}(v)$ for all $v \epsilon \mathrm{P} \cap [\mathrm{H}, \mathrm{H}]$. Let m be the least common multiple of the denominators of the rational numbers $\hat{\alpha}(r_j)$ for some finite set of free generators r_j of P, and let $\tau(r) = m\hat{\sigma}(r) - m\hat{\alpha}(r)$ for $r \epsilon \mathrm{P}$. By Lemma 2 again, $\hat{\tau}(r)$ may be extended to a homomorphism of H into the rationals, which we shall also denote by $\hat{\tau}$. The functions $\tau_t(z) = m\sigma_t(z) - \hat{\tau}(t)$ have the same values as $m\sigma_t(z)$ whenever $t \epsilon \mathrm{P} \cap [\mathrm{H}, \mathrm{H}]$, and $\tau_t(z) \epsilon Z$ whenever $t \epsilon \mathrm{P}$. Thus $\nu_T(z) = \exp \tau_t(z)$ defines a factor of automorphy with the character class $\chi_1(\nu) = m\chi_g$.

We shall call a group Γ *integral* if every form χ_g itself represents the character class of a factor of automorphy. The examples to be considered in Chapter 4 will clearly be seen to be integral. It should be noted here that, upon applying the decomposition theorems for the operators ∂ and $\bar{\partial}$ in the obvious manner, $\bar{\partial}\partial g(z) = \psi(z) + \bar{\partial}\partial h(z)$ where $\psi(z)$ is a harmonic differential form and $h(z)$ is a Γ-invariant C^∞ function on D. Therefore for the purposes of Theorem 5 it is sufficient to consider only those functions $g(z)$ for which $\bar{\partial}\partial g(z)$ is a harmonic differential form.

Now assuming that the group Γ contains no transformations with fixed points, the differential forms $\phi(z) = \bar{\partial}\partial g(z)$ of Theorem 5 represent in the de Rham sense the character class $\chi_1(\nu)$ considered as a 2-cocycle of D/Γ. Theorem 4 may be expressed more analytically in terms of this differential form.

COROLLARY *to Theorem* 4. *If* $\phi(z) = \sum_{j,k} \phi_{j\bar{k}}(z) dz^j \wedge d\bar{z}^k$ *is a differential form representing the character class* $\chi_1(\nu)$ *of a positive factor of automorphy, then the Hermitian matrix* $(-i\phi_{j\bar{k}}(z))$ *is positive definite in the mean on harmonic forms, in the following sense: for any non-trivial closed differential form*

$$\theta(z) = \sum_j (-1)^j \theta_j(z) dz^1 \wedge \cdots \wedge dz^{j-1} \wedge dz^{j+1} \wedge \cdots \wedge dz^n,$$

$$-i \int_{D/\Gamma} \sum_{j,k} \phi_{j\bar{k}}(z) \theta_j(z) \bar{\theta}_k(z) dv > 0,$$

where dv *is the positive volume element on the manifold* D/Γ. (Γ is assumed to contain no transformations with fixed points.)

Proof. Theorem 4 asserts that

$$i^{(1-n^2)} (\chi_1(\nu) \cup \theta \cup \bar{\theta}) [D/\Gamma] > 0,$$

where θ is the cohomology class represented by the differential form $\theta(z)$. However in terms of differential forms

$$(\chi_1(\nu) \cup \theta \cup \bar{\theta}) [D/\Gamma] = \int_{D/\Gamma} \phi(z) \wedge \theta(z) \wedge \bar{\theta}(z)$$

$$= 2^n i^{(n^2-2)} \int_{D/\Gamma} \sum_{j,k} \phi_{j\bar{k}}(z) \theta_j(z) \bar{\theta}_k(z) dx^1 \wedge \cdots \wedge dx^{2n},$$

where $z_j = x_{2j-1} + ix_{2j}$, from which the assertion follows.

9. It is perhaps appropriate to discuss at this point the connection between factors of automorphy as considered here and the related concept of

a complex line bundle [14]. If Γ has no fixed points, then every factor of automorphy defines a complex line bundle on the manifold D/Γ in the obvious manner; the presence of fixed points means not only that D/Γ need not be a complex manifold, but also that the line bundle defined may be locally multiple-valued. To proceed in the opposite direction, any complex line bundle on D/Γ, when D/Γ is a complex manifold, induces a complex line bundle on D. If D is a Stein manifold, for example [5, 6], then all topologically trivial line bundles on D are also analytically trivial, and are thus equivalent to line bundles induced by factors of automorphy; in particular if $H^2(D, R) = 0$ for a Stein manifold D, all line bundles on D/Γ are equivalent to bundles induced by factors of automorphy.

In the present note, factors of automorphy have been treated on their own, without utilizing their relationship to complex line bundles. This method has the advantages that the concepts and constructions can be developed in terms of the group Γ and its action on the manifold D, and that they express in a natural manner the global nature and multiplicity at fixed points which distinguish factors of automorphy from line bundles. In addition, the forms in which factors of automorphy appear, and such results as Theorem 4, are of more interest from a function-theoretic point of view than from the point of view of line bundles.

IV. Examples.

10. Our first example illustrates a method for constructing non-trivial factors of automorphy in fairly general cases. Letting $\{\omega_\alpha(z)\}$ be a basis for the complex linear space of abelian differentials on D/Γ and $\{w_\alpha(z)\}$ be the corresponding abelian integrals, introduce the functions $g(z) = \sum_{\alpha,\beta} \xi_{\alpha\bar\beta} w_\alpha(z) \bar{w}_\beta(z)$ for arbitrary complex constants $\xi_{\alpha\bar\beta}$. The differential forms

$$\bar\partial\partial g(z) = -\sum_{\alpha,\beta} \xi_{\alpha\bar\beta}\omega_\alpha(z) \wedge \bar\omega_\beta(z) \tag{10}$$

are obviously Γ-invariant, so that they define elements

$$\chi_g \in \operatorname{Hom}(\mathrm{P} \cap [\mathrm{H}, \mathrm{H}]/[\mathrm{P}, \mathrm{H}]; C^+)$$

as in Section 8. If χ_g is integral valued for some choice of $\xi_{\alpha\bar\beta}$, Theorem 5 asserts that a multiple of χ_g represents the character class of a factor of automorphy. The proof of Theorem 5 was actually constructive in nature; examining that proof, one sees that the factors of automorphy so represented are given explicitly by

$$\nu_T(z) = \exp\left[\sum_{\alpha\beta} \xi_{\alpha\bar\beta}\bar\omega_\beta(T) w_\alpha(z) + \zeta_T\right], \tag{11}$$

or by integral powers of these functions if necessary; the ζ_T are suitably chosen complex constants and $\omega_\beta(T)$ are the periods of the abelian integrals $w_\beta(z)$. Factors of automorphy of the form (11) may be called generalized theta factors, and the associated relatively automorphic functions generalized theta functions [18].

It is perhaps of some interest to examine the classical theta functions [19, 20] from this point of view. For this purpose, let D be the entire n-dimensional complex affine space and Γ be a group of translations generated by the translations along $2n$ real linearly independent vectors in D. The abelian differentials are then simply the differential forms dz^α, $\alpha=1,\cdots,n$, and the periods of the associated abelian integrals corresponding to a transformation $T\in\Gamma$ are the components of the vector representing the translation T. A differential form $\phi=\sum_{\alpha,\bar\beta}\phi_{\alpha\bar\beta}(z)dz^\alpha\wedge d\bar z^\beta$ is harmonic and Γ-invariant if and only if the coefficients $\phi_{\alpha\bar\beta}(z)$ are harmonic, Γ-invariant functions on D, hence constants; that is to say, all harmonic Γ-invariant differential forms of type $(1,1)$ are of the form (10). Therefore all factors of automorphy are equivalent to factors (11), which in this case have the even simpler form

$$\nu_T(z)=\exp[\sum_{\alpha,\bar\beta}\xi_{\alpha\bar\beta}\bar\omega^\beta(T)z_\alpha+\zeta_T]. \tag{12}$$

The latter statement was first proved in more than one complex variable by Appell [1].

The character class has a well-known representation in this case. Letting $T_1,\cdots,T_{2n}$ be the translations generating the group Γ, we may write $\Gamma\cong\mathrm{H}/\mathrm{P}$ where H is a free group on corresponding generators $t_1,\cdots,t_{2n}$ and P is the normal subgroup of H generated by the words $v_{jk}=t_jt_kt_j^{-1}t_k^{-1}$. Since $\mathrm{P}\cap[\mathrm{H},\mathrm{H}]=\mathrm{P}$, the group Γ is certainly integral. The character class $\chi_1(\nu)$ of the factor (12) is determined by the values

$$\chi_1(\nu)[v_{jk}]=\frac{1}{2\pi i}\sum_{\alpha,\bar\beta}\xi_{\alpha\bar\beta}[-\omega_{\alpha j}\bar\omega_{\beta k}+\omega_{\alpha k}\bar\omega_{\beta j}].$$

In matrix form $\Omega=(\omega_{\alpha j})$ is the period matrix, either in the classical sense or in the sense of Section 2, $\Xi=(\xi_{\alpha\bar\beta})$, and $\mathrm{X}_1(\nu)=(\chi_1(\nu)[v_{jk}])$ is given by

$$\mathrm{X}_1(\nu)=\frac{1}{2\pi i}[{}^t\bar\Omega{}^t\Xi\Omega-{}^t\Omega\Xi\bar\Omega]. \tag{13}$$

The differential form $\sum_{\alpha,\bar\beta}\xi_{\alpha\bar\beta}dz^\alpha\wedge d\bar z^\beta$ therefore represents the character class of a factor of automorphy if and only if the matrix (13) is integral; furthermore a group Γ with period matrix Ω will admit a non-trivial factor of auto-

morphy if and only if there exists a constant matrix Ξ such that (13) is integral.

The above is one part of the condition that Ω be a Riemann matrix. If we require that the group Γ with Ω as period matrix admit a positive factor of automorphy, we secure, as with Frobenius [9], the complete condition that Ω be a Riemann matrix. For, referring to the corollary to Theorem 4 and considering the closed differential forms $\theta(z)$ where $\theta_j(z)$ are constants, we derive immediately that $-i\Xi$ is a positive definite Hermitian matrix. It is in this sense that we may refer to Theorem 4 as a generalization of the theorem of Frobenius. The converse implication would be of considerable interest, but nothing so strong is yet known.

The more general factors (11) have also been used previously. P. Myrberg has considered such factors in one complex variable in his studies of the analytic representation of automorphic functions [17]. In this connection, it will follow in the next section that whenever a factor (11) in one complex variable has a non-trivial character class, then all factors of automorphy are equivalent to rational multiples of that factor; hence the Myrberg factors are equivalent, insofar as representing automorphic functions abstractly, to the Poincaré factors which we shall consider next.

11. If D is a bounded subdomain of the complex affine space and Γ is a properly discontinuous group of analytc homeomorphisms of D onto itself, we may introduce the Poincaré factors of automorphy $\{J_T(z)\}$, where $J_T(z)$ denotes the complex Jacobian determinant of the mapping T. A differential form representing the character class of this factor in the sense of Theorem 5 can be constructed immediately. Considering T as a mapping on a real $2n$-dimensional Euclidean space, its real Jacobian determinant is $j_T(z) = |J_T(z)|^2$. If $ds^2 = \sum_{j,k} g_{jk}dx^j dx^k$ is a Γ-invariant real metric on D and $g(z) = \det(g_{jk}(z))$, then $[g(Tz)]^{\frac{1}{2}} = j_T(z)[g(z)]^{\frac{1}{2}}$; hence, as in Theorem 5, the desired differential form is just

$$\phi(z) = \bar{\partial}\partial \log[g(z)]^{\frac{1}{2}} = \sum_{\alpha,\beta} R_{\alpha\bar{\beta}} dz^\alpha \wedge d\bar{z}^\beta, \tag{14}$$

where $R_{\alpha\bar{\beta}}$ is the complex Ricci curvature tensor. Consequently the necessary and sufficient condition that all factors of automorphy be equivalent to rational powers of the Poincaré factor is that all Γ-invariant differential forms $\bar{\partial}\partial g(z)$ with rational periods on the cycles $\Phi(v)$ of Section 1 be cohomologous to rational multiples of the complex Ricci curvature form (13). Although this condition is certainly not always fulfilled. for example when D

is the Cartesian product of domains D_1 and D_2, and Γ is the direct product of a group Γ_1 on D_1 and Γ_2 on D_2, nevertheless it is fulfilled trivially in one complex variable. Therefore in one complex variable all factors of automorphy are equivalent to rational powers of the Poincaré factors $J_T(z) = dT(z)/dz$.

The results in one complex variable can also be obtained directly from the known structural properties of the group. We may select a set of free generators $s_1, \cdots, s_p, t_1, \cdots, t_p, u_1, \cdots, u_q$ for H such that P is the normal subgroup generated by the words $u_1^{m_1}, \cdots, u_q^{m_q}$,

$$w = u_q \cdots u_1 s_p^{-1} t_p^{-1} s_p t_p \cdots s_1^{-1} t_1^{-1} s_1 t_1.$$

Then $\mathrm{P} \cap [\mathrm{H}, \mathrm{H}]/[\mathrm{P}, \mathrm{H}]$ is the infinite cyclic group generated by the coset of $w^m (u_1^{-m_1})^{m/m_1} \cdots (u_q^{-m_q})^{m/m_q}$, where $m = \text{l.c.m.}(m_1, \cdots, m_q)$. The character class is therefore completely determined by its value on the generator of the group $\mathrm{P} \cap [\mathrm{H}, \mathrm{H}]/[\mathrm{P}, \mathrm{H}]$, which may be called the characteristic number. If a factor of automorphy ν has a non-trivial character class, so that its characteristic number $M \neq 0$, then it follows from Theorem 2 that all factors of automorphy are equivalent to the factors $\nu_T(z)^{m/M}$ for integers m. In particular, since it follows from (13) that the Poincaré factors have non-trivial character class, all factors are equivalent to the factors $(dT(z)/dz)^{m/M}$. If no fixed points are present, it is well known that $M = 2p - 2$, where p is the genus of the Riemann surface D/Γ.

PRINCETON UNIVERSITY.

REFERENCES.

[1] P. Appell, "Sur les fonctions périodiques de deux variables," *Journal de Mathématiques Pures et Appliquées*, vol. 97 (1884), pp. 16-48.

[2] W. L. Baily, "On the quotient of an analytic manifold by a group of analytic homeomorphisms," *Proceedings of the National Academy of Sciences*, vol. 40 (1954), pp. 804-808.

[3] S. Bochner, "On compact complex manifolds," *Journal of the Indian Mathematical Society*, vol. 11 (1947), pp. 1-21.

[4] ———, and W. T. Martin, *Several Complex Variables*, Princeton, 1948.

[5] H. Cartan, "Variétés analytiques complexes et cohomologie," *Colloque sur les Fonctions de Plusieurs Variables*, Brussels, 1953.

[6] ———, *Séminaire École Normale Supérieure*, 1951-1952 (mimeographed notes,

reprinted by the Department of Mathematics, Massachusetts Institute of Technology).

[7] S. Eilenberg and S. MacLane, "Group extensions and homology," *Annals of Mathematics*, vol. 43 (1942), pp. 757-831.

[8] ———, "Cohomology theory in abstract groups," *Annals of Mathematics*, vol. 48 (1947), pp. 51-78.

[9] G. Frobenius, "Ueber die Grundlagen der Theorie der Jacobischen Funktionen," *Journal für die reine und angewandte Mathematik*, vol. 97 (1884), pp. 16-48.

[10] P. R. Garabedian and D. C. Spencer, "A complex tensor calculus for Kähler manifolds," *Acta Mathematica*, vol. 89 (1953), pp. 279-331.

[11] R. C. Gunning, "General factors of automorphy," *Proceedings of the National Academy of Sciences*, vol. 44 (1955), pp. 496-498.

[12] W. V. D. Hodge, *The Theory and Application of Harmonic Integrals*, Cambridge, 1952.

[13] H. Hopf, "Fundamentalgruppe und zweite Bettische Gruppe," *Commentarii Mathematici Helvetici*, vol. 14 (1941), pp. 257-309.

[14] K. Kodaira and D. C. Spencer, "Groups of complex line bundles over compact Kahler varieties," *Proceedings of the National Academy of Sciences*, vol. 39 (1953), pp. 868-872.

[15] B. O. Koopman and A. B. Brown, "On the covering of analytic loci by complexes," *Transactions of the American Mathematical Society*, vol. 34 (1932), pp. 231-251.

[16] S. Lefschetz and J. H. C. Whitehead, *Transactions of the American Mathematical Society*, vol. 34 (1933), pp. 510-528.

[17] P. Myrberg, "Sur les fonctions automorphes," *Annales Scientifiques de l'École Normale Supérieure*, vol. 68 (1951), pp. 383-424.

[18] G. deRham and K. Kodaira, *Harmonic Integrals* (mimeographed notes, Institute for Advanced Study, 1950).

[19] C. L. Siegel, *Analytic Functions of Several Complex Variables* (mimeographed notes, Institute for Advanced Study, 1949).

[20] A. Weil, "Théorèmes fondamentaux de la théorie des fonctions thêta," *Seminaire Bourbaki*, May, 1949.

RATIONAL EQUIVALENCE OF ARBITRARY CYCLES.* [1]

By Pierre Samuel.

We intend to give a definition and some properties of the notion of rational equivalence for cycles of arbitrary dimension on a non singular projective variety. That there exists such a theory of rational equivalence for arbitrary cycles was demonstrated by the work of F. Severi on the "series of equivalence." In this paper we will not attempt to make explicitly the connection with Severi's theory, but the influence of his work will be easy to detect. We will content ourselves to provide the working geometer with a certain number of tools. The definition of rational equivalence that we give is analogous to the definition of algebraic equivalence given by A. Weil ([6]), and the principal results we prove are also valid for algebraic equivalence, with trivial modifications in the proofs; it may even be observed that the only result in which we use non-elementary methods (i. e. the specialization theorem, in the proof of which we use both the degeneration principle and a property of the divisors of the second kind) has, in the case of algebraic equivalence, an analogue which is trivial. It goes without saying that many proofs have been inspired by those in A. Weil's paper ([6]). I am very grateful to J. I. Igusa, G. Washnitzer and O. Zariski for their valuable encouragement and their invaluable advice during the preparation of this paper.

1. Preliminary results. We use the terminology and notations of A. Weil ([5]), sometimes modified according to a recent book of ours ([4]). As we shall work only with *non singular* varieties, it will be sufficient, in order to prove that the intersection product $X \cdot Y$ of two cycles on a variety V is defined, to check that all the components of $\mathrm{Supp}(X) \cap \mathrm{Supp}(Y)$ [2] have the right dimension. We first recall some well known facts, which will be useful in the sequel:

(a) In order to apply the associativity formula to the intersection of, let us say, three cycles X, Y, Z on a variety V, it is sufficient to show that

* Received September 14, 1955.

[1] This work was supported by a research project at Harvard University, sponsored by the Office of Ordnance Research, United States Army, under contract DA-19-020-ORD-3100.

[2] By $\mathrm{Supp}(X)$ we mean the reunion of the components of the cycle X.

the intersection product obtained by "associating" X, Y, Z in a given order are defined.

(b) For the projection formula $Y \cdot \mathrm{pr}_U X = \mathrm{pr}_U((Y \times V) \cdot X)$ on the product of two non singular projective varieties to be valid, it is sufficient that its right hand side (i. e. $(Y \times V) \cdot X$) be defined.

Given a homogeneous cycle X on a variety V, it is convenient to consider its *codimension*, i. e. the integer $\dim(V) - \dim(X)$; we denote it by $\mathrm{cod}_V(X)$, or $\mathrm{cod}(X)$ when no confusion may arise. Then, if the intersection product $X \cdot Y$ of two cycles X, Y on V is defined, we have

(c) $\mathrm{cod}_V(X \cdot Y) = \mathrm{cod}_V(X) + \mathrm{cod}_V(Y)$.

If W is a subvariety of V such that $X \cdot W$ is defined, we have

(d) $\mathrm{cod}_W(X \cdot W) = \mathrm{cod}_V(X)$.

As in [4], Chap. I, § 10, No. 3, we do not restrict the notion of *rational mapping* of V into W to those mappings F such that $F(V) = W$. We identify the rational mapping F with its graph in $V \times W$; the fact that F is a rational mapping means that F is a variety, that $\mathrm{pr}_V(F) = V$, and that the projection index of F in V is equal to 1. We recall that, for a rational mapping F of V into W to be *regular* (i. e. regular at every point of V), it is necessary and sufficient that the birational correspondence between F and V defined by pr_V be *biregular* ([4]). It is easily seen that, if F is a rational mapping of V into W and if Y is a cycle on W such that $F^{-1}(Y) = \mathrm{pr}_V((V \times Y) \cdot F)$ is defined, then we have

(e) $\mathrm{cod}_V(F^{-1}(Y)) = \mathrm{cod}_W(Y)$

(cf. [4], Chap. II, § 6, No. 9, e)). It can also be proved, by using the same method as in Chap. II, § 6, No. 9, f), that, if F is a *regular* mapping of V into W, and if Y and Y' are two cycles on W such that $F^{-1}(Y)$, $F^{-1}(Y')$, $Y \cdot Y'$ and $F^{-1}(Y \cdot Y')$ are defined, then $F^{-1}(Y) \cdot F^{-1}(Y')$ is defined, and we have

(f) $F^{-1}(Y \cdot Y') = F^{-1}(Y) \cdot F^{-1}(Y')$.

It may be observed by using (a) that, at least in the case in which V and W are non singular, it is sufficient to assume that $Y \cdot Y'$ and $F^{-1}(Y \cdot Y')$ are defined, i. e. that $(V \times Y) \cdot (V \times Y') \cdot F$ is defined.

LEMMA 1. *Let S, R and V be three non singular projective varieties, Z a cycle in $V \times R$ and F a regular mapping of S into R. Then the mapping F' of $V \times S$ into $V \times R$ defined by $F'(v \times s) = v \times F(s)$ is regular. If a*

is a point of S such that $Z(F(a)) = \mathrm{pr}_V((V \times F(a)) \cdot Z)$ is defined, then the cycles $Z' = F'^{-1}(Z)$ on $V \times S$ and $Z'(a)$ in V are defined, and we have $Z'(a) = Z(F(a))$.

Proof. The regularity of F' is straightforward. As we shall work in the quadruple product $V \times S \times V \times R$, we shall, in order to avoid confusions, denote its third factor by V': $V \times S \times V' \times R$. If we denote by V^D the diagonal of the product $V \times V'$, the graph of F' is $F \times V^D (F \subset R \times S$, $V^D \subset V \times V')$. Our hypothesis that $Z(F(a)) = \mathrm{pr}_{V'}((V' \times F(a)) \cdot Z)$ is defined means that $(V' \times F(a)) \cdot Z$ is defined (intersection in $V' \times R$). We first prove that the intersection cycle

$$T = (V \times S \times Z) \cdot (V \times a \times V' \times R) \cdot (F \times V \times V') \cdot (S \times R \times V^D)$$

is defined. Since F is regular, $(a \times R) \cdot F$ is defined and equal to $a \times F(a)$; whence the intersection product of the two middle factors of T is defined, and is equal to $V \times a \times V' \times F(a)$. We are thus reduced to proving that

$$T = (V \times S \times Z) \cdot (V \times a \times V' \times F(a)) \cdot (S \times R \times V^D)$$

is defined. Since $(V' \times F(a)) \cdot Z$ is defined and equal to $Z(F(a)) \times F(a)$, the intersection product of the two first factors in T is defined and equal to $V \times a \times Z(F(a)) \times F(a)$. We are now reduced to prove that

$$T = (V \times a \times Z(F(a)) \times F(a)) \cdot (S \times R \times V^D)$$

is defined. Since this cycle, in $S \times R \times V \times V'$, is

$$(a \times F(a)) \times (V^D \cdot (V \times Z(F(a)))),$$

it is obviously defined, and its projection on V is $Z(F(a))$.

By the remark (a) above, all the partial intersection cycles in the formula for T are defined. In particular the intersection product of the first, third and fourth factors is defined; since it is equal to $(V \times S \times Z) \cdot (F \times V^D)$ $= (V \times S \times Z) \cdot F'$, this shows that $Z' = F'^{-1}(Z) = \mathrm{pr}_{V \times S}((V \times S \times Z) \cdot F')$ is defined. On the other hand we have, by the projection formula

$$\mathrm{pr}_{V \times S}(T) = \mathrm{pr}_{V \times S}(((V \times S \times Z) \cdot F') \cdot (V \times a \times V' \times R)) = (V \times a) \cdot Z'.$$

This proves that $(V \times a) \cdot Z'$ is defined, whence also $Z'(a) = \mathrm{pr}_V((V \times a) \cdot Z')$. We therefore have $Z'(a) = \mathrm{pr}_V(\mathrm{pr}_{V \times S}(T))$, whence $Z'(a) = \mathrm{pr}_V(T) = Z(F(a))$.

LEMMA 2. *Let Z^z be a subvariety of the product $V^v \times W^w$ of two non singular varieties V, W, and let A^a be a subvariety of V. Denote by V_j the set of all points P of V such that $\dim(Z \cap (P \times W)) \geqq j$. Then V_j is a*

closed subset of V. If A intersects properly all the components of all the closed sets V_j, then the intersection product $Z\cdot(A\times W)$ is defined (whence also $Z(A)=\mathrm{pr}_W(Z\cdot(A\times W))$).

Proof. For every integer j such that $V_j\neq V_{j+1}$ we denote by A_j the point set $(V_j-V_{j+1})\cap A$. The fact that V_j is a closed set follows from [4], Chap. I, § 8, No. 2, c). Then A_j is a finite union of open subvarieties of V, and, by hypothesis, each of them has a dimension $\leqq a-v+\dim(V_j)$. Since we have $\dim(Z(P))=j$ for every P in V_j-V_{j+1}, the principle of counting constants ([4], Chap. I, § 10, No. 2) shows that the dimension of $Z\cap(A_j\times W)$ is $\dim(A_j)+j$ whence at most $a-v+\dim(V_j)+j$. We see in the same way that $\dim(Z\cap(V_j\times W))=\dim(W_j)+j$, whence $\dim(W_j)+j\leqq\dim(Z)=z$ since $Z\cap(V_j\times W)\subset Z$. It follows that we have $\dim(Z\cap(A_j\times W))\leqq a-v+z$. Since A is the union of the sets A_j (which are in finite number), $Z\cap(A\times W)$ is the union of the sets $Z\cap(A_j\times W)$, whence we have the inequality $\dim(Z\cap(A\times W))\leqq a-v+z=z+(a+w)-(v+w)$. Since $z+(a+w)-(v+w)$ is the proper dimension for $\dim(Z\cap(A\times W))$, this proves our assertion as $V\times W$ is non singular.

LEMMA 3. *Let V^n be a non singular projective variety imbedded in P_q, A^a a subvariety of V, (B_j) a finite family of subvarieties of V. Then, for almost every linear variety L^{q-n-1} the projecting cone C of A with vertex L is such that, if we write $C\cdot V=A+R$, then the residual intersection R properly intersects all the varieties B_j.*

Proof. If $\dim(B_j)+a\geqq n$, let B_j' be a generic plane section of B_j of dimension $n-a-1$; then R intersects properly B_j if and only if $R\cap B_j$ is empty. In other words, replacing B_j by B_j', we may assume that we have $\dim(B_j)+a<n$, and have to prove that L may be chosen in such a way that $R\cap B_j$ is empty for every j. Let W_j be the reunion of all the straight lines joining a point of A and a point of B_j (and of all their specializations; if (a) and $(b^{(j)})$ are affine generic points of A and B_j over a common field of definition k, and if t is a trancendental element over $k(a,b)$, then W_j is the locus of $(ta+(1-t)b^{(j)})$ over k). The dimension of W_j is $\leqq a+\dim(B_j)+1\leqq n$. Thus almost all linear varieties L^{q-n-1} have an empty intersection with all the W_j's. For such a linear variety L, and for every point (a) of A, the linear variety L'^{q-n} containing L and (a) has at most (a) as common point with B_j. More precisely, if (b) is a point of $B_j\cap R$, then the linear variety L'^{q-n} determined by (b) and L intersects A at a point (a) which must coincide with (b), otherwise the line $(a)(b)$ would

meet L, in contradiction with the choice of L. Now, since the point $(a) = (b)$ is common to A and R, since $C \cdot V = A + R$, and since $(a) = (b)$ is a simple point of V, the tangent linear variety of V at $(a) = (b)$ must contain the projecting linear variety L'; then, since L' is tangent to V at a point common to A and B_j, it must contain a specialization of the line joining a generic point of A to a generic point of B_j, i.e., a line lying on the variety W_j; this again contradicts the choice of L. Therefore $B_j \cap R$ is empty.

2. Definition and characterizations of rational equivalence.

Definition. *Let V^n be a non singular projective variety. A cycle X^r on V^n is said to be rationally equivalent to* 0 *if there exist a non singular unirational variety R^m, a cycle Z^{m+r} on $V \times R$ and two points a and b of R^m such that $Z(a) = \mathrm{pr}_V(Z \cdot (V \times a))$ and $Z(b)$ are defined and that $X = Z(a) - Z(b)$.*

We recall that a variety R is said to be *unirational* if its absolute function field is a subfield of a purely transcendental extension of the universal domain. We denote by $\mathfrak{R}_r(V)$ the set of all r-cycles on V which are rationally equivalent to 0.

Theorem 1. *The set of cycles $\mathfrak{R}_r(V)$ is a group under addition.*

Proof. Let X and X' be two elements of $\mathfrak{R}_r(V)$. We write

$$X = Z(a) - Z(b), \qquad X' = Z'(a') - Z'(b')$$

where a, b (resp. a', b') are points of a non singular unirational variety R (resp. R'), and where Z (resp. Z') is a cycle on $V \times R$ (resp. $V \times R'$). We consider, on $V \times R \times R'$ the cycle $U = Z \times R' - R \times Z'$. Since

$$\begin{aligned} U(a \times a') &= \mathrm{pr}_V(U \cdot (V \times a \times a')) \\ &= \mathrm{pr}_V((Z \cdot (V \times a)) \times a' - (Z' \cdot (V \times a')) \times a) = Z(a) - Z'(a'), \end{aligned}$$

and since, similarly, $U(b \times b') = Z(b) - Z'(b')$, we have

$$X - X' = U(a \times a') - U(b \times b').$$

As $R \times R'$ is a unirational variety, this proves that $X - X' \in \mathfrak{R}_r(V)$.

Theorem 2. *If an r-cycle X on a non singular projective variety V^n is rationally equivalent to* 0 *there exist two points a, b of the projective line P_1 and a positive cycle T on $V \times P_1$ such that $T(a)$ and $T(b)$ are defined and that $X = T(a) - T(b)$.*

Proof. We write $X = Z(a) - Z(b)$, Z, a, b being as in the definition. We first reduce ourselves to the case in which Z is positive. We write $Z = Z^+ - Z^-$, where Z^+ and Z^- are the positive and the negative part of Z (cf. [4], Chap. I, § 9, No. 2, c), and we consider, on $R \times V \times R$, the cycle $U = Z^+ \times R + R \times Z^-$. By a simple computation as in Theorem 1, we see that

$$U(a \times b) = \mathrm{pr}_V(U \cdot (a \times V \times b)) = Z^+(a) + Z^-(b)$$

(these cycles being defined since $Z(a)$ and $Z(b)$ are defined). Similarly we have $U(b \times a) = Z^+(b) + Z^-(a)$, whence $X = U(a \times b) - U(b \times a)$. Since $R \times R$ is a unirational variety, we have achieved the reduction to the case of a positive cycle U.

If we now show that any two points of a unirational variety may be connected by a rational curve, the proof of Theorem 2 will be complete. In fact we shall have a rational mapping F of P_1 into $R \times R$ and two points c and d of P_1 such that $F(c) = a \times b$ and $F(d) = b \times a$. Since P_1 is a non singular curve, the local rings of all its points are valuation rings, and this proves that F is regular. Our conclusion then follows from Lemma 1, Section 1. We are thus reduced to proving the following lemma:

LEMMA 4. *Given two points a, b of a unirational variety V, there exists a rational curve lying on V and joining a and b.*

Proof of the lemma. The following terminology will be convenient. We say that a variety W *dominates* a variety U at a point u of U if there exists a rational mapping F of W onto U and a point w of W such that F is regular at w and that $F(w) = u$; we say that W dominates U if it dominates U at every point of U. According to Lüroth's theorem we may replace V by any rational variety which dominates V at a and b. Since there exists a rational mapping H of a projective space P_q onto V, we first replace V by the graph H, and we may thus assume the existence of a birational and regular mapping F of V onto P_q. As a second step we show the existence of a birational correspondence T between P_q and another projective space P_q' such that P_q' dominates V at a, and that P_q' corresponds biregularly to P_q at $F(b)$; let a' be a point of P_q' having a as regular image on V. If we apply the same result to P_q', the join V' of V and P_q', a' and $(b, T^{-1}(F(b)))$ instead of P_q, V, $F(b)$ and a, we obtain a projective space P_q'' which dominates V at a and b. Since any two points of a projective space may be joined by a straight line, this will prove the lemma.

Therefore we need only to show the existence of a birational correspondence T between P_q and P_q' such that P_q' dominates V at a and that P_q' corresponds biregularly to P_q at $F(b)$. We are immediately reduced to the affine case, where a, b, $F(a)$, $F(b)$ are at finite distance. Let $x_1, \cdots, x_q$ be the coordinates in the affine space A_q, and $(x_1, \cdots, x_q, z_1, \cdots, z_q)$ a generic point of V over an algebraically closed field k; the quantities z_i are rational functions of $x_1, \cdots, x_q$ over k. We may assume that the points a and b are rational over k, and that $F(a)$ and a are the origins in A_q and A_{q+r}.

We shall apply the classical method of *successive quadration transformations* (analogous results are proved in the work of O. Zariski, by which our proof is directly inspired). We first notice the existence of a discrete valuation of rank 1 of $k(x_1, \cdots, x_q)$ which is zero-dimensional (i.e., since k is algebraically closed, the corresponding place ϕ takes its values in k), and which admits a as center on V. The existence of such an "analytical arc" is well known (see O. Zariski, "Foundations of a general theory of birational correspondences," *Transactions of the American Mathematical Society,* vol. 53 (1943), pp. 490-542; this statement is proved on pp. 501-502 as case (a) in Theorem 5; the method of proof shows that the constructed valuation may be assumed to be discrete). The numbers $v(x_j)$, $v(z_i)$ are then > 0. After a suitable linear change of coordinates we may assume that $F(b)$ does not lie on $X_1 = 0$, that $v(x_1)$ is the smallest of the numbers $v(x_j)$, and (replacing x_j by $x_j - \phi(x_j/x_1)x_1$) that $v(x_j) > v(x_1)$ for $j \geqq 2$.

The quadratic transformation $x_1 = x_1'$, $x_j = x_1'x_j'$ for $j \geqq 2$ is then biregular at $F(b)$. For any polynomial $D(x)$ we have $D(x) = (x_1')^d D'(x')$, where d is the order of $D(x)$ (i.e. the degree of its lowest degree form), and where $D'(x')$ is a polynomial uniquely determined by $D(x)$. If d is $\neq 0$, we have $v(D'(x')) < v(D(x))$ and $v(x_1') \leqq v(D(x))$. We say that a finite family of polynomials is adequate if each rational function z_i is a power product of these polynomials. Let $(A_1(x), \cdots, A_s(x))$ be such a family. Then z_i, considered as a rational function of the variables (x'), is a power product of the polynomials x_1', $A_u'(x')$, and the family $(x_1', A_1'(x'), \cdots, A_s'(x'))$ is adequate (for the variables (x')). If one at least of the orders $v(A_u(x))$ is $> v(x_1)$, we have the inequality

$$\max(v(x_1'), v(A_u'(x'))) < \max(v(A_u(x))). \tag{I}$$

Since v is a discrete valuation of rank 1, we cannot repeat this procedure of quadratic transformations an infinite number of times and always get an inequality of the type (I) between the orders (for v) of the elements of two successive adequate families. Therefore we eventually get a system $(y_1, \cdots, y_q)$

of independent variables (with $v(y_j) > 0$ for every j), and an adequate family $(B_u(y))$ such that either $v(B_u(y)) = 0$ (in which case $B_u(y)$ is a polynomial with non-zero constant term), or $v(B_u(y)) = v(y_1)$. By a last quadratic transformation $y_1 = y_1'$, $y_j = y_1 y_j'$ for $j \geqq 2$ (we assume as above that $v(y_j) > v(y_1)$), we get, for the indices u such that $v(B_u(y)) = v(y_1)$, $B_u(y) = y_1' B_u'(y')$, and $B_u(y) = B_u'(y')$ for the others; at any rate all the polynomials $B_u'(y')$ have a constant term $\neq 0$. Hence we can write $z_i = (y_1')^{m(i)} R_i(y')$, where $R_i(y')$ is a rational function belonging to the local ring $\mathfrak{o}$ of the origin in the affine space with coordinates (y'). On the other hand the elements x_j are polynomials in the variables (y') by construction. Since the numbers $v(z_i)$, $v(y_1')$ are > 0, we have $m(i) > 0$ for every i. Thus all the elements z_i, x_j belong to the local ring $\mathfrak{o}$.

THEOREM 3. *Let X^r be an r-cycle on a non singular projective variety V^n. For X to be rationally equivalent to* 0 *on V it is necessary and sufficient that there exist a Chow variety W of positive r-cycles on V, a rational curve C on W and two points a, b of C such that $X = c(a) - c(b)$ ($c(x)$ denoting the cycle corresponding to the Chow point x).*

Proof. We first prove the necessity of our condition. Since X is rationally equivalent to 0, we can write $X = Z(c) - Z(d)$, where Z is a positive cycle on $V \times P_1$, and where c, d are points of P_1 (Theorem 2). Let t be a generic point of P_1 over an algebraically closed field of definition k of c, d, V and Z. Since $Z(a)$ and $Z(b)$ are specializations of $Z(t)$ over k ([4], Chap. II, § 6, No. 8, a), the Chow points ([4], Chap. I, § 9, No. 6) belong to a common Chow variety W of positive cycles (and W is defined over k; cf. [4], Chap. I, § 9, No. 6). Since the Chow point x of $Z(t)$ is rational over $k(Z(t)) \subset k(t)$, the locus C of x over k is a rational curve by Lüroth's theorem. Taking for a and b the Chow points of the cycles $Z(c)$ and $Z(d)$, we immediately see that a and b lie on C. Thus the necessity is proved.

We now prove the sufficiency. *Let* $X = c(a) - c(b)$, where a and b denote two points of a rational curve C lying on a Chow variety W of positive cycles on V. Let k be an algebraically closed field of definition of V, W, C, a, b, and let x be a generic point of C over k. The cycle $c(x)$ is rational over some purely inseparable extension K of $k(x)$ ([4], Chap. I, § 9, No. 4, g). As k is perfect and as $k(x)$ is a simple transcendental extension of k, say $k(t')$ (the system (t') being reduced to one element), there exists an exponent e such that $K = k(t'^{p^{-e}})$. Thus K is a simple transcendental extension $k(t)$ $(t = t'^{p^{-e}})$. Since the cycle $c(x)$ is rational over $k(t)$, we may consider its

locus Z in $V \times P_1$ over k ([4], Chap. II, § 6, No. 8, b): $\mathrm{pr}_V((V \times t) \cdot Z) = c(x)$. If F denotes the regular mapping of $V \times P_1$ onto $V \times C$ such that $F(v \times t) = x \times c(x)$, the point set Z is the inverse image $F^{-1}(T)$ of the incidence correspondence T attached to the system C of cycles ([4], Chap. I, § 10, No. 4, b) (as to cycles we have $F^{-1}(T) = p^f \cdot Z$). We denote by c and d two points of P_1 which are mapped into a and b by the restriction of F to P_1. Since $T(a)$ and $T(b)$ (considered as point sets) have the right dimension and are equal, as cycles, to $p^f \cdot c(a)$ and $p^f \cdot c(b)$, Lemma 1, Section 1 shows that the cycles $p^f \cdot Z(c)$ and $p^f \cdot Z(d)$ are defined and equal to $p^f \cdot c(a)$ and $p^f \cdot c(b)$, whence $X = Z(c) - Z(d)$. Therefore the sufficiency is proved since P_1 is a unirational variety.

We notice that, if a *divisor* X^{n-1} on V^n is rationally equivalent to 0, then a theorem proved in [5] (Chap. VIII, No. 2, Th. 5) shows that it is linearly equivalent to 0. The converse is obvious.

We say that two r-cycles X, X' on a non singular projective variety V are *rationally equivalent* on V if their difference $X - X'$ is rationally equivalent to 0 (i. e. if $X - X' \in \mathfrak{R}_r(V)$). Since $\mathfrak{R}_r(V)$ is a group (Theorem 1), this is actually an equivalence relation. Since, in the case of divisors, it coincides with linear equivalence, we may denote it by the same symbol, and write $X \sim X'$.

3. Some properties of rational equivalence.

THEOREM 4 ("*specialization theorem*"). *Let X^r be an r-cycle on a projective non singular variety V, which is rationally equivalent to* 0 *on V. If X' is a specialization of X over some field of definition k of V, then X' is rationally equivalent to* 0 *on V.*

Proof. Since every specialization over k is the same thing as a specialization over the algebraic closure of k followed by a k-automorphism, and since everything algebraic is preserved by k-automorphisms, we may assume that k is algebraically closed. By Theorem 3 there exists a Chow variety W of r-cycles on V, a rational curve C on W, and two points a and b of C such that $X = c(a) - c(b)$. We extend the specialization $X \to X'$ to a k-specialization $(X, C, a, b) \to (X', C', a', b')$. Since W is defined over k and is complete, C' is a 1-cycle on W, and a', b' are two points of C'. If we show that C' is connected and that all its components are rational curves, the proof of Theorem 4 will be complete: in fact we have $X' = c(a') - c(b')$, and there exist rational curves $C_1', \cdots, C_n'$ on W and connecting points $d_1, \cdots, d_{n-1}$ such that $a', d_1 \in C_1', d_1, d_2 \in C_2', \cdots, d_{n-1}, b' \in C_n'$; then the cycles $c(a') - c(d_1)$,

$c(d_1)-c(d_2),\cdots,c(d_{n-1})-c(d')$ are rationally equivalent to 0 by Theorem 3, whence also their sum X' by Theorem 1. We are thus reduced to proving:

LEMMA 5. *If a 1-cycle C' (in some projective space P_n) is a specialization over k of a rational curve C, then C' is connected, and all its components are rational curves.*

That C' is connected follows from the degeneration principle ([8]). Replacing C by a generic specialization $\bar{C}$ over k, and denoting by $(\bar{c})$ and (c') the Chow points of $\bar{C}$ and C', Prop. 7 in App. II of [5] shows that there exists an algebraically closed field K containing k, a set of quantities (x), a non singular curve U defined over K and admitting $(x,\bar{c})$ as generic point over K, and a point (x',c') of U: in fact we may suppose that (c') is a non generic specialization of (c) (otherwise our assertion is trivial), and the facts that we may take K algebraically closed and U non singular (i. e. normal over K) follow from the proof of the above quoted Prop. 7. Since $\bar{C}$ is irreducible, all the coefficients of the components of $\bar{C}$ are prime to the characteristic (there is only one of them, which is equal to 1), and a result of Chow (cf. [7], App., Prop. 6) shows that $\bar{C}$ is a rational curve over $k(\bar{c})$; whence $K(x,\bar{c})$ is a field of definition of the curve $\bar{C}$. Let (y) be a generic point of $\bar{C}$ over $K(x,\bar{c})$. Let S $(\subset P_q\times U)$ be the surface locus of $(y,\bar{c},x)$ over K. Its horizontal section $(P_q\times(c',x'))\cap S$ is equal to C', at least as a point set.

Since $\bar{C}$ is a rational curve, $K(x,\bar{c})(y)$ is a function field of genus 0 over $K(x,\bar{c})$. Since K is algebraically closed, and since $\dim_K(K(x,\bar{c}))=1$, $K(x,\bar{c})$ is quasi algebraically closed by a theorem of Tsen. Thus $K(x,\bar{c})(y)$ is a simple transcendental extension $K(x,\bar{c})(t)$ of $K(x,\bar{c})$, since a field of genus 0 admits a positive rational divisor of degree 2, and since a conic over a quasi-algebraically closed field admits a rational point. In other words the surface S is birationally equivalent to the product $P_1\times U$ (this is a well known result of Max Noether): more precisely there exists a birational correspondence T between S and $P_1\times U$ such that $T(x,\bar{c},y)=(x,\bar{c},t)$.

We have to prove that every component D of $(P_q\times(c',x'))\cap S$ is a rational curve. It is clear that every subvariety of $P_1\times U$ which corresponds to D under T lies on the straight line $P_1\times(x',c')$. We consider a prime divisor v of the function field of S (over K) having D as a center on S. If the center of v on $P_1\times U$ is the entire line $P_1\times(x',c')$, then the residue field R_v of the valuation v is the function field of this line, whence is a purely transcendental extension of K (the point (x',c') is rational over K since K is algebraically closed, and since it is non-generic specialization of $(x,\bar{c})$). Otherwise the center of v on $P_1\times U$ is a simple point of $P_1\times U$ (then the

prime divisor v is of the second kind on $P_1 \times U$), and a result about divisors of the second kind on surfaces ([1]) shows that the residue field R_v is also a purely transcendental extension of K in this case. Since R_v is the function field of some curve D^0 corresponding to D in a derived normal model S^0 of S, this proves that D^0 is a rational curve. Since D is a "projection" of D^0, D is also a rational curve by Lüroth's theorem. Lemma 5 and Theorem 4 are thereby proved.

THEOREM 5. *Let X and Y be two cycles of dimension r and s on a non singular projective variety V^n. If $X \sim 0$ and if $X \cdot Y$ is defined, then $X \cdot Y \sim 0$.*

Proof. We write $X = Z(a) - Z(b)$, where a and b are simple points of a rational variety R, and where Z is a cycle on $V \times R$ such that $Z(a)$ and $Z(b)$ are defined. We first study a particular case:

LEMMA 6. *If $Z(a) \cdot Y$ and $Z(b) \cdot Y$ are defined, then $X \cdot Y$ is rationally equivalent to 0 on V. If, furthermore, Y is a non singular subvariety of V, then $X \cdot Y$ is rationally equivalent to 0 on Y.*

In fact, if $Z(a) \cdot Y = Y \cdot \mathrm{pr}_V((V \times a) \cdot Z)$ is defined, then the set $(V \times a) \cap \mathrm{Supp}(Z) \cap \mathrm{Supp}(Y \times R)$, which is in biregular correspondence with $\mathrm{Supp}(Y) \cap \mathrm{Supp}(Z(a))$ has the correct dimension, whence (Section 1, (a)) $(V \times a) \cdot Z \cdot (Y \times R)$ and $Z' = Z \cdot (Y \times R)$ are defined. By the projection formula we have $Z(a) \cdot Y = \mathrm{pr}_V((V \times a) \cdot Z \cdot (Y \times R)) = Z'(a)$, whence $X = Z'(a) - Z'(b)$, and this proves our first assertion. For the second one we consider Z' as a cycle on $Y \times R$, and we still have $Z'(a) = \mathrm{pr}_Y((Y \times a) \cdot Z')$. This proves the lemma.

For completing the proof of Theorem 5, we take a field of definition k of all the components of V, R, a, b, X, Y and Z, and we consider Y as a specialization over k of a cycle $\bar{Y}$ such that $\bar{Y} \cdot Z(a)$ and $\bar{Y} \cdot Z(b)$ are defined: for example we take a projecting cone C of Y whose vertex is a linear variety L^{q-n-1} generic over k (q: dimension of a projective space in which V^n is imbedded), a generic projective transform $\bar{C}$ of C (over $k(C)$), and set $\bar{Y} = \bar{C} \cdot V + D$, where D is the residual intersection $C \cdot V - Y$ (Section 1, Lemma 3). Then $\bar{Y} \cdot X$ is rationally equivalent to 0 by Lemma 6. Since we are in a case where the specialization theorem ([4], Chap. II, § 6, No. 7) may be applied separately to positive and negative parts of the cycles we consider, $Y \cdot X$ is a specialization of $\bar{Y} \cdot X$, whence is ~ 0 on V by Theorem 4.

We shall prove later that the second assertion of Lemma 5, is still true even if $Z(a) \cdot Y$ and $Z(b) \cdot Y$ are not defined ($X \cdot Y$ being, of course, defined).

COROLLARY 1. *The group $\mathfrak{R}_r(V)$ contains the intersection cycles $D_1 \cdot D_2 \cdot \cdots \cdot D_{n-r}$ which are defined and such that one of the divisors D_i is ~ 0.*

COROLLARY 2. *For an r-cycle X in projective space P_n to be rationally equivalent to 0 in P_n, it is necessary and sufficient that it be of degree 0.*

In fact we use a theorem of Severi ([4], Chap. II, § 6, No. 4) and we write X as an intersection $X = D_1 \cdot \cdots \cdot D_{n-r}$ of divisors. If X has degree 0, then one of the D's must have degree 0 by Bezout's theorem; whence one of the D_i's is ~ 0, and also X by Theorem 5. For the converse we write $X = Z(a) - Z(b)$, and notice that $Z(a)$ and $Z(b)$ have the same degree.

COROLLARY 3. *A cycle X on V which is a complete intersection (i.e. $X = V \cdot Y$, Y being a cycle in the ambiant projective space P_q) is rationally equivalent to 0 on V if and only if it is of degree 0 (as a cycle in P_q).*

In fact the part "only if" is clear. Conversely, if X has degree 0, then Y has also degree 0 by Bezout's theorem. By Corollary 2 we may write $Y = D_1 \cdot \cdots \cdot D_s$ where the D_i's are divisors in P_q such that D_1 is of degree 0. Then the $X' = D_1^+ \cdot D_2 \cdot \cdots \cdot D_s \cdot V$, $X'' = D_1^- \cdot D_2 \cdot \cdots \cdot D_s \cdot V$ are defined, and since $D_1 = D_1^+ - D_1^-$ is a representation of D_1 under the form $Z(a) - Z(b)$, the relation $X = X' - X''$ is a representation of X under the form $Z'(a) - Z'(b)$. The conclusion follows by Lemma 5.

THEOREM 6. *Let V, W be two non singular projective varieties, and X a cycle on V which is ~ 0 on V. Then $X \times W \sim 0$ on $V \times W$.*

Proof. We write $X = Z(a) - Z(b)$ (a, b simple points of unirational variety R, Z cycle in $V \times R$). We set $Z' = Z \times W$ (in $V \times W \times R$). Then, by the formula on intersections on product varieties ([4], Chap. II, § 6, No. 5, f), $Z'(a)$ is defined and equal to $Z(a) \times W$. Thus $X \times W = Z'(a) - Z'(b)$.

COROLLARY. *If X is ~ 0 on V and if Y is any cycle on W, then $X \times Y \sim 0$ on W.*

Analogous proof by using $Z'' = Z \times Y$. Or notice that

$$X \times Y = (X \times W) \cdot (V \times Y),$$

and use Theorems 6 and 5.

THEOREM 7. *Let V, W be two non singular projective varieties, and let X be a cycle on $V \times W$. If $X \sim 0$ on $V \times W$, then $\mathrm{pr}_V(X) \sim 0$ on V.*

Proof. We write $X = Z(a) - Z(b)$ (a, b simple points of unirational variety R, Z cycle on $V \times W \times R$). We have

$$\mathrm{pr}_V(Z(a)) = \mathrm{pr}_V(\mathrm{pr}_{V\times W}(V \times W \times a) \cdot Z)$$
$$= \mathrm{pr}_V(\mathrm{pr}_{V\times R}((V \times W \times a) \cdot Z) = \mathrm{pr}_V((V \times a) \cdot \mathrm{pr}_{V\times R}(Z))$$

(we apply the projection formula, as recalled in (b), § 1). Thus, if we set $Z' = \mathrm{pr} V \times R(Z)$, we have $\mathrm{pr}_V(X) = Z'(a) - Z'(b)$.

THEOREM 8. *Let V and W be two non singular projective varieties, F a rational mapping of V into V, and X a cycle on V which is ~ 0 on V. Then, if $F^{-1}(X)$ is defined, it is ~ 0 on W.*

Proof. We have $F^{-1}(X) = \mathrm{pr}_W((W \times X) \cdot F)$, where $(W \times X) \cdot F$ is defined. Since $X \sim 0$, we have $W \times X \sim 0$ by Theorem 6, whence $(W \times X) \cdot F \sim 0$ on $W \times V$ by Theorem 5. Therefore $\mathrm{pr}_W((W \times X) \cdot F) \sim 0$ on W by Theorem 7.

It may be observed that, if Z is any cycle on $W \times V$, and X a cycle on V which is ~ 0 on V, then $Z(X) = \mathrm{pr}_W((W \times X) \cdot Z)$ is ~ 0 on W if it is defined. The proof is the same as in Theorem 8.

COROLLARY. *Let W be a non singular subvariety of a non singular projective variety V. If X is a cycle on V which is ~ 0 on V, and if $W \cdot X$ is defined, then $W \cdot X$ is ~ 0 on W.*

In fact, if we denote by i the inclusion mapping of W into V, we have $W \cdot X = i^{-1}(X)$.

Remark. It is *not* true that, if Y is a cycle on W ($\subset V$) which is ~ 0 *on* V, then it is ~ 0 *on* W: take $V = P_3$, W a quadric, Y the difference of two straight lines of different systems on W. But the converse is true: if Y is a cycle on W ($\subset V$) which is ~ 0 *on* W, then Y is ~ 0 on the bigger variety V. This follows from Theorem 3 and from the fact that a Chow variety of positive r-cycles on W is a subvariety of a Chow variety of r-cycles on V; or one may apply the remark following Theorem 8 to $i(Y)$.

We may use Lemma 5 (about specializations of rational curves) to prove a more general result than Theorem 4:

THEOREM 9. *Let X^r be an r-cycle on a non singular projective V^n, and let (X', V') be a specialization of (X, V) over some field k, such that the cycle V' is a non singular variety. If $X \sim 0$ on V, then $X' \sim 0$ on V'.*

Proof. It is well know that the relation "$\mathrm{Supp}(A) \subset \mathrm{Supp}(B)$" between two cycles A, B may be expressed by a system of equations with absolutely algebraic coefficients in the Chow coordinates of A and B ([2]; or [4], Chap. I, § 9, No. 7, g where an analogous result is proved). Thus X' is a cycle on V'. We now use Theorem 3 and write $X = c(a) - c(b)$ where a, b are points of a Chow variety W of positive r-cycles on V, such that a and b can be connected by a rational curve C lying on W. We extend the given specialization to $(V, X, W, C, a, b) \to (V', X', W', C', a', b')$. By [4], Chap. I, § 9, No. 7, h every point of $\mathrm{Supp}(W')$ is a specialization of a point of W; thus every point of $\mathrm{Supp}(W')$ is the Chow point of some positive r-cycle on V'; whence also a', b', and every point of $\mathrm{Supp}(C')$. Since C' is a connected union of rational curves (Lemma 5), it follows, as in the proof of Theorem 4, that $c(a') - c(b')$ is ~ 0 on V'. Since the relation $X = c(a) - c(b)$ gives, by specialization, $X' = c(a') - c(b')$, we have $X' \sim 0$ on V'.

THEOREM 10. *Let V^n be a non singular variety in projective space P_q. If a cycle X of dimension r is ~ 0 on V, then there exists a function f on V and a cycle Y^{r+1} on V such that $X = (f) \cdot Y$.*

Proof. We first prove that X may be written as $X = \sum_i (f_i) \cdot Y_i$. We write $X = Z(a) - Z(b)$, with $Z \subset V \times P_1$, $a, b \in P_1$. By decomposition of the cycle Z in irreducible components, we are reduced to the case in which Z is irreducible. Then $\mathrm{pr}_V(Z)$ is an irreducible subvariety W of V, whose dimension is r or $r+1$. If $\dim(W) = r$, then $Z(a) = Z(b) = W$, $X = 0$, and there is nothing to prove. Otherwise $Z(a)$ and $Z(b)$ are divisors on W, and, since their Chow-points may be connected by a rational curve of "W-divisors $Z(t)$," they are linearly equivalent *on* W. Denoting by H a suitable divisor of degree 0 in P_q, we thus have

$$X = Z(a) - Z(b) = (H \cdot W)_{P_q} = ((H \cdot V)_{P_q} \cdot W)_V.$$

Since $(H \cdot V)_{P_q}$ is the divisor of a function f on V, the first part of the proof is complete.

We may now write $X = \sum_i (f_i) \cdot Y_i^{r+1}$ where the Y_i's are distinct subvarieties of V, whence $X = \sum_i H_i \cdot Y_i$, where the H_i's are divisors of degree 0 in P_q. Let $P_i(x) = 0$ be the equation of H_i, $P_i(x)$ being the quotient of two forms of like degree. We can find forms $F_j(x)$ of the same degree such that F_j is $\neq 0$ on Y_j but is 0 on Y_i for $i \neq j$. Then the divisor H with equation $F(x) = (\sum_j F_j(x))^{-1} (\sum_j F_j(x) P_j(x)) = 0$ has degree 0 and is such

that $H\cdot Y_j = H_j\cdot Y_j$ for every j: in fact F and F_j induce the same function on Y_j. If f_j denotes the function induced by F on V, we have

$$X = \sum_j (f)\cdot Y_j = (f)\cdot \sum_j Y_j.$$

We conclude this section by showing that our notion of rational equivalence coincides with the notion of "linear equivalence of cycles of arbitrary dimension" defined by A. Weil in [5], p. 959 (bottom). We use here the word "rational" instead of "linear" since it has been used by F. Severi in a long series of papers for denoting a notion which is also the same as our notion, whereas A. Weil devotes only two lines to the "linear" equivalence of cycles of arbitrary dimension.

First it is clear that our groups $\mathfrak{R}_r(V)$ satisfy conditions (A), (B), (C′), (D), (E) and (L) of Weil. These properties are respectively stated in Theorem 6, Cor. to Theorem 8, Theorem 7, Theorem 5, Theorem 7, except for (L) which is evident. Thus our groups $\mathfrak{R}_r(V)$ contain the corresponding groups of Weil.

Conversely, if a cycle X on V is rationally equivalent to 0 (in our sense), we have $X = Z(a) - Z(b) = \mathrm{pr}_V(Z\cdot(V\times((a)-(b))))$, where a and b are two points of P_1 and Z a cycle on $V\times P_1$. In the sense of Weil the cycles $(a)-(b)$, $V\times((a)-(b))$, $Z\cdot(V\times((a)-(b)))$ and X are "linearly" equivalent to 0" by (L), (A), (D) and (E) respectively. This proves the opposite inclusion.

4. The ring of rational equivalence classes.

THEOREM 11. *Let V be a non singular projective variety, and let α and β be two rational equivalence classes of cycles on V. There exist cycles A and B in α and β such that $A\cdot B$ is defined. And the rational equivalence class of $A\cdot B$ depends only on α and β.*

Proof. We choose A and B' arbitrarily in α and β. If $A\cdot B'$ is not defined, we consider a common field of definition k of V, A, B', a generic linear variety L^{q-n-1} (q: dimension of the ambiant projective space) over k, the projecting cone C of B' with vertex L, and a generic projective transform $\bar{C}$ of C over $k(C)$. We set $C\cdot V = B' + D$, and $\bar{C}\cdot V = E$. The cycle $A\cdot D$ is defined by Lemma 3, § 1, and $A\cdot E$ is evidently also defined. We now take $B = E - D$; then $A\cdot B$ is defined. Since the group of projective transformations is a rational variety, we have $C \sim \bar{C}$ on P_q, whence $C\cdot V \sim \bar{C}\cdot V$ on V (Cor. to Theorem 8). In other words we have $B' + D \sim E$ on V, i.e., $B' \sim B$. This proves our first assertion.

Let now A, A' be elements of α and B, B' be elements of β such that $A \cdot B$ and $A' \cdot B'$ are defined. The first part of the proof shows the existence of B'' in β such that $A \cdot B''$ and $A' \cdot B''$ are defined. Then Theorem 5 shows that $A \cdot B \sim A \cdot B''$, that $A \cdot B'' \sim A' \cdot B''$, and that $A' \cdot B'' \sim A' \cdot B'$. Whence $A \cdot B \sim A' \cdot B'$ by transitivity. Therefore the rational equivalence class of $A \cdot B$ does not depend on the choice of A and B in α and β.

It follows from Theorem 11 that we have defined, on the set $\mathfrak{C}(V)$ of rational equivalence classes of cycles on V, a *multiplication*. The commutativity and the associativity of intersection products show that this multiplication is commutative and associative. We have also, on $\mathfrak{C}(V)$, an addition, defined by the addition of cycles; this addition is distributive with respect to the multiplication. The set $\mathfrak{C}_r(V)$ of rational equivalence classes of cycles of *codimension* r is an additive subgroup of $\mathfrak{C}(V)$, $\mathfrak{C}(V)$ is the direct sum of the $\mathfrak{C}_r(V)$, and, by (c), § 1, the product of an element of $\mathfrak{C}_r(V)$ and of an element of $\mathfrak{C}_s(V)$ is in $\mathfrak{C}_{r+s}(V)$. In other words $\mathfrak{C}(V)$, graded by the codimension, becomes a *graded commutative ring*; we call it the *ring of rational equivalence classes* on V. It admits a unit element, the equivalence class of V itself.

It is easily seen that the rational equivalence classes of the cycles which are *algebraically equivalent* to 0 on V form an *ideal* $\mathfrak{A}$ in $\mathfrak{C}(V)$. The factor ring $\mathfrak{C}(V)/\mathfrak{A}$ is canonically isomorphic to the ring of algebraic equivalence classes on V (defined by a method similar to the one used here for rational equivalence).

Examples. 1) If V is the projective space P_n, then the graded ring $\mathfrak{C}(V)$ is isomorphic to $Z[X]/(X^{n+1})$, the coset of X corresponding to the class of the hyperplanes (cf. Cor. 2 to Theorem 5).

2) If V is a quadric in P_3, $\mathfrak{C}(V)$ is isomorphic to $Z[X,Y]/(X^2,Y^2)$, the cosets of X and Y corresponding to the classes of the straight lines on V.

Given two non singular projective varieties V and W, and a *correspondence* Z between V and W (i.e. a cycle on $V \times W$), we are going to define an *additive mapping* Z^* of $\mathfrak{C}(W)$ into $\mathfrak{C}(V)$. In fact, given a rational equivalence class α on W, there exists a cycle A in α such that

$$Z(A) = \mathrm{pr}_V((V \times A) \cdot Z)$$

is defined: in fact the proof of Theorem 11 shows that, given a finite number of subvarieties (B_j) of W, there exists A in α such that $A \cdot B_j$ is defined for every j; thus our assertion follows from Lemma 2, § 1. On the other hand,

if A and A' are elements of α such that $Z(A)$ and $Z(A')$ are defined, then we have $Z(A) \sim Z(A')$ on V, by the remark following Theorem 8.

The mapping Z^* is *not* a ring homomorphism in general. However, if F is a *regular mapping* of V into W, we may use the formula $F^{-1}(A \cdot B) = F^{-1}(A) \cdot F^{-1}(B)$ recalled in § 1, (f). More precisely, given two rational equivalence classes α and β on V, we first choose A in α such that $F^{-1}(A)$ is defined (by Lemma 2, § 1 this amounts to requiring that A intersects properly a finite number of subvarieties W_j of W). We have now to choose B in β in such a way that $A \cdot B$, $F^{-1}(B)$ and $F^{-1}(A \cdot B)$ are defined; this means that B must intersect properly all the components of A, all the subvarieties W_j and all the components of $\mathrm{Supp}(A) \cap W_j$ for every j; this again is a requirement of proper intersection with a finite number of subvarieties of V, and this requirement may be fulfilled as in the proof of Theorem 11. Therefore the additive mapping F of $\mathfrak{C}(W)$ into $\mathfrak{C}(V)$ defined by F^{-1} satisfies $F^*(\alpha \cdot \beta) = F^*(\alpha) \cdot F^*(\beta)$ for every α, β. In other words F^* is a *homomorphism* of the ring $\mathfrak{C}(W)$ into $\mathfrak{C}(V)$, and this homomorphism preserves codimensions by § 1 (e), i. e. is *a homomorphism for the structures of graded rings of* $\mathfrak{C}(W)$ and $\mathfrak{C}(V)$.

If U, V, W are three non singular projective varieties, and $F: U \to V$ and $G: V \to W$ are *regular mappings,* then the composition $G \circ F$ is a regular mapping of U into W. It follows immediately from [4], Chap. II, § 6, No. 9, g that, when they are defined, the cycles $F^{-1}(G^{-1}(A))$ and $(G \circ F)^{-1}(A)$ (A: cycle on W) are equal. Consequently we have

$$F^*(G^*(\alpha)) = (G \circ F)^*(\alpha) \text{ for every } \alpha \text{ in } \mathfrak{C}(W).$$

In other words the homomorphisms F^* are *transitive.* We can also say that $\mathfrak{C}$ is a *contravariant functor* for the categories of non singular projective varieties and of regular mappings.

UNIVERSITE DE CLERMONT-FERRAND AND
HARVARD UNIVERSITY.

REFERENCES.

[1] S. Abhyankar, "On the valuations centered in a local domain," (to appear).

[2] W.-L. Chow and B. L. van der Waerden, "Uber zugeordnete Formen und algebraische systeme von algebraischen mannigfalligkeiten," *Mathematische Annalen,* vol. 113 (1937), pp. 692-704.

[3] T. Matsusaka, "The theorem of Bertini on linear systems in modular fields," *Memoirs College Science Kyoto University*, vol. 26 (1950), pp. 51-62.

[4] P. Samuel, "Méthodes d'Algèbre Abstraite en géométrie Algébrique," *Ergebnisse der Mathematik und ihrer Grenzgebiete*, Neue Folge, Heft 4 (1955), Berlin (Springer).

[5] A. Weil, *Foundations of Algebraic Geometry*, American Mathematical Society Colloquium Publications, 1946.

[6] ———, "Sur les critères d'équivalence en géométrie algébrique," *Mathematische Annalen*, vol. 128 (1954), pp. 95-127.

[7] ———, "On algebraic groups of transformations," *American Journal of Mathematics*, vol. 77 (1955), pp. 355-391.

[8] O. Zariski, "Holomorphic functions," *Memoirs of the American Mathematical Society*, No. 5 (1951).

[9] ———, "Pencils on an algebraic variety and a new proof of the theorem of Bertini," *Transactions of the American Mathematical Society*, vol. 50 (1941), pp. 48-70.

SOME BASIC THEOREMS ON ALGEBRAIC GROUPS.*

By Maxwell Rosenlicht.[1]

The subject of algebraic groups has had a rapid development in recent years. Leaving aside the late research by many people on the Albanese and Picard variety, it has received much substance and impetus from the work of Severi on commutative algebraic groups over the complex number field, that of Kolchin, Chevalley, and Borel on algebraic groups of matrices, and especially Weil's research on abelian varieties and algebraic transformation spaces. The main purpose of the present paper is to give a more or less systematic account of a large part of what is now known about general algebraic groups, which may be abelian varieties, algebraic groups of matrices, or actually of neither of these types.

The first two parts of our work are devoted largely to extending, by very similar methods, the work of Nakano and Weil on the construction of transformation spaces, homogeneous spaces, and factor groups. Our third part proves the expected homomorphism theorems, the fourth gives a useful result on the existence of cross sections, and the last part gives the principal structure theorems.

The main result of this paper, Theorem 16, was announced by Chevalley in 1953, together with a proof whose basic idea was to consider the natural homomorphism from a connected algebraic group to its Albanese variety and then apply the basic properties of Albanese and Picard varieties. Chevalley's theorem also appears in the recent publications of I. Barsotti (*Ann. Mat. Pura Appl.*, vol. 38 (1955) and *Atti Accad. Naz. Lincei Rend. Cl. Sci. Fis. Mat. Nat.*, ser. 8, vol. 18 (1955)); his papers represent work done approximately simultaneously with the author's and seem to follow a similar method. We have made no use of Barsotti's papers except to appropriate from him the statement of the first half of our Proposition 2, which is not needed elsewhere.

An older version of this paper was completed before the author knew of the existence of Weil's recent papers [4], [5]; it contained roughly the

* Received December 14, 1955.

[1] During the final revision of this paper the author was connected with a project of the U. S. Air Force.

same results we present now, but in much weaker form. The main differences with the present paper are due to the fact that our original statement of Theorem 1 (which was proved by the method of Weil and Nakano) assumed that the group G contained a dense set of rational points.

Following N. Bourbaki, a map $\tau\colon V \to W$ will be called *surjective* if $\tau(V) = W$. Otherwise, the terminology and conventions we employ are for the most part those of Weil ([3], [4], [5], [6]). In particular, we follow his systematic use in [4] and [5] of the language of the Zariski topology, even for bunches of varieties. However we remark one important divergence with his usage: in this paper the closed subsets of a variety V are taken to include V itself.

The author is indebted to A. Weil, whose advice in the preparation of the final manuscript made possible many simplifications and generalizations.

1. Algebraic groups as transformation groups. We define an *algebraic group* to be the union G of a finite number of disjoint algebraic varieties $\{G_\alpha\}$ (called the *components* of G) together with a group structure on G such that for any components G_α, G_β of G, the map $g_1 \times g_2 \to g_1 g_2^{-1}$ restricted to $G_\alpha \times G_\beta$ is an everywhere defined rational map of $G_\alpha \times G_\beta$ into some component G_γ of G. A field k is called a *field of definition* of G if it is a field of definition for each component of G and for all the above rational maps $G_\alpha \times G_\beta \to G_\gamma$. If G has only one component, we say that G is *connected.*

If G is an algebraic group having k as a field of definition, consideration of the map $g \to gg^{-1} = e$ shows that the identity e of G is rational over k. It follows that the map $g \to g^{-1}$ is an everywhere defined rational map on each component of G, that the map $g_1 \times g_2 \to g_1 g_2$ is an everywhere defined rational map on each pair of components of G, and that these two maps are defined over k. If G_0 is the component of G that contains e, then $G_0 G_0^{-1} = G_0$, so G_0 is a connected algebraic group having k as a field of definition. If $g \, \varepsilon \, G$ and G_α is the component of G that contains g, then $e \, \varepsilon \, g^{-1} G_\alpha$, so $x \to g^{-1}x$ is a rational map of G_α into G_0. Similarly, $y \to gy$ is a rational map of G_0 into G_α, and it follows that G_0 and G_α are biregularly birationally equivalent, with $G_\alpha = gG_0$. Similarly $G_\alpha = G_0 g$. Thus G_0 is a normal subgroup of G and each component of G is a coset of G_0. Since G_0 is nonsingular, so is any component of G. Finally, if $g_1, g_2 \, \varepsilon \, G$, we have $k(g_1, g_2) = k(g_1, g_1 g_2) = k(g_2, g_1 g_2)$.

Let G be an algebraic group and V a variety. We say that G *operates regularly* on V (in the terminology of Weil, V is a *transformation space*

for G) if for each component G_α of G we are given an everywhere defined rational map $g \times v \to g(v)$ of $G_\alpha \times V \to V$ such that

(1) $g_1(g_2(v)) = g_1g_2(v)$ for any $g_1, g_2 \varepsilon G$, $v \varepsilon V$.

(2) $e(v) = v$ for any $v \varepsilon V$.

In this case it is clear that if k is a field of definition for G, V and the operation of G on V, and if $g \varepsilon G$, $v \varepsilon V$, then $k(g, g(v)) = k(g, v)$.

Let G be an algebraic group and V a variety. We say that G *operates* on V (or that V is a *pre-transformation space for* G) if for each component G_α of G we are given a rational map $g \times v \to g(v)$ of $G_\alpha \times V \to V$ such that if k is a field of definition for G, V, and each of these rational maps and if $g_1 \times g_2 \times v$ is a generic point over k of $G_\alpha \times G_\beta \times V$ (G_α, G_β being any components of G) then

(1) $g_1(g_2(v)) = g_1g_2(v)$.

(2) $k(g_1, g_1(v)) = k(g_1, v)$.

Consideration of the graphs of the various rational maps $G_\alpha \times V \to V$ shows this definition to be independent of the choice of the field of definition k. If G operates regularly on V, then G operates on V. When there is no danger of confusion, we shall usually write gv instead of $g(v)$. If $g_1(g_2v)$ and $(g_1g_2)v$ are both defined, then they must be equal, so we may simply write g_1g_2v.

LEMMA. *Let the algebraic group G operate on the variety V and let k be a field of definition for G, V, and the operation of G on V. If $a, b \varepsilon G$, and x is a simple point of V such that both bx and $a(bx)$ are defined, then $(ab)x$ is defined and equals $a(bx)$. If v is generic for V over $k(a)$, then av is defined and $k(a, av) = k(a, v)$; in particular av is generic for V over $k(a)$. If v is generic for V over k, then $ev = v$.*

Let g_1, g_2 be independent generic points over k of the components of G that contain a, b respectively, and let v be generic for V over $k(g_1, g_2)$. Then g_1g_2 is generic over k for the component of G that contains ab and v is generic for V over $k(g_1g_2)$. Let the point y of an affine representative of V that contains $a(bx)$ be such that $ab \times x \times y$ is a specialization over k of $g_1g_2 \times v \times (g_1g_2)v$. Then $a \times ab \times x \times y$ is a specialization over k of $g_1 \times g_1g_2 \times v \times (g_1g_2)v$. If we extend this latter specialization to a specialization over k of $g_1 \times g_2 \times g_1g_2 \times v \times g_2v \times g_1(g_2v) \times (g_1g_2)v$ we get $a \times b \times ab \times x \times bx \times a(bx) \times y$. Hence $y = a(bx)$ is uniquely determined by ab and x. It follows that $(ab)x$ is defined and equals $a(bx)$. For the rest,

we may assume that a is rational over k. If g is generic for a component of G over $k(v)$, the same is true of ga, so $(ga)v$ is defined and $k(g,(ga)v) = k(ga,(ga)v) = k(ga,v) = k(g,v)$. Thus $(ga)v$ is generic for V over $k(g) = k(g^{-1})$, so $g^{-1}((ga)v)$ is defined and generic for V over $k(g)$; thus av is defined and generic for V over $k(g)$. Since $k(g,(ga)v) = k(g,v)$, v is rational over $k(g,av)$, hence over $k(av)$; thus $k(av) = k(v)$. Finally, to show that $ev = v$, it suffices to show that $e(ev) = ev$, and this is known.

THEOREM 1. *Let the algebraic group G operate on the variety V and let k be a field of definition for G, V, and the operation of G on V. Then there exists a variety V', birationally equivalent over k to V, such that the operation of G on V' that is induced by its operation on V is regular.*

If G is connected, this is part of the main theorem of [4]. In the general case, note that G_0 operates on V, so that we may suppose that G_0 operates regularly on V. Let $\gamma_1, \cdots, \gamma_r$ be a set of points of G that are algebraic over k and such that each component of G contains at least one γ_i and let W be a k-closed proper subset of V such that $\gamma_1 v, \cdots, \gamma_r v$ are all defined whenever $v \varepsilon V - W$. If $a \varepsilon G_0$ and $v \varepsilon V - W$ then $a(\gamma_i v)$ is defined, so $(a\gamma_i)v$ is also defined, whence gv is defined for any $g \varepsilon G$ and $v \varepsilon V - W$. If $a \varepsilon G_0$, $g \varepsilon G$, $v \varepsilon V - aW$, then $a^{-1}v$ is defined and not on W, so $ga(a^{-1}v)$ is defined, whence gv is defined. Hence gv is defined for all $g \varepsilon G$, $v \not\varepsilon W' = \bigcap_a aW$, a ranging over all points of G_0. W' is clearly k-closed and $aW' = W'$ for any $a \varepsilon G_0$. Let W'' be the set of all points $p \varepsilon V$ for which there exists a $g \varepsilon G$ such that gp is defined and on W'. If $p \varepsilon W''$ we can choose $a \varepsilon G_0$ and $i = 1, \cdots, r$ such that $(a\gamma_i)p$ is defined and on W', whence $\gamma_i p$ is defined and on W'. Hence W'' is a closed proper subset of V which contains W'. Since W'' is invariant with respect to all k-automorphisms of the universal domain, W'' is k-closed. If $p \varepsilon V - W''$ and $g \varepsilon G$ then clearly $gp \not\varepsilon W''$. Hence the theorem is proved by taking $V' = V - W''$.

The following corollary give a kind of uniqueness result for Theorem 1.

COROLLARY. *Let V_1, V_2 be birationally equivalent varieties, let G be an algebraic group which operates regularly on both V_1 and V_2 in a manner consistent with their birational equivalence, and let k be a field of definition for G, V_1, V_2, the birational equivalence between V_1 and V_2, and the operation of G on V_1 and V_2. Then there exist k-closed proper subsets F_1, F_2 of V_1, V_2 respectively such that G operates regularly on both $V_1 - F_1$ and $V_2 - F_2$ and such that the birational correspondence between $V_1 - F_1$ and $V_2 - F_2$ is biregular.*

For there exist k-closed proper subsets W_1, W_2 of V_1, V_2 respectively such that the birational correspondence between $V_1 - W_1$ and $V_2 - W_2$ is biregular (for example, cf. [1, § 2, Lemma 1]) and it suffices to set

$$F_1 = \bigcap_{g \varepsilon G} gW_1, \qquad F_2 = \bigcap_{g \varepsilon G} gW_2.$$

As an example of the corollary, if our V_1, V_2 are homogeneous spaces with respect to G (i.e., G operates regularly and transitively on V_1 and V_2), then F_1 and F_2 are empty.

Let V, W be varieties and $\tau\colon V \to W$ a rational map. If k is a field of definition for V, W, τ and v is a generic point of V over k, we say that τ is *generically surjective* if τv is generic for W over k, that τ is *separable* if $k(v)$ is separably generated over $k(\tau v)$, and that τ is *purely inseparable* if $k(v)$ is a purely inseparable algebraic extension of $k(\tau v)$. Consideration of the graph of τ on $V \times W$ shows these definitions to be independent of the field of definition k.

In the following discussion leading up to Theorem 2, G will denote an algebraic group that operates on the variety V and k will denote a field of definition for G, V, and the operation of G on V. If $g \varepsilon G$ and f is a rational function on V we define another rational function $\lambda_g f$ on V by the rule $\lambda_g f(p) = f(g^{-1}p)$. Since $\lambda_{g_1}\lambda_{g_2} = \lambda_{g_1 g_2}$, the map $g \to \lambda_g$ is a homomorphism of G into the group of automorphisms of the function field on V that leave constant functions fixed. If $g \varepsilon G$ and K is an extension field of k over which g is rational, then λ_g defines an automorphism of the field $K(V)$ of all rational functions on V that are defined over K, for if v is any generic point of V over K, then $K(V)$ is naturally isomorphic to $K(v)$ and we have $K(g^{-1}v) = K(v)$.

If f is a rational function on V, we say that f is *invariant* if $\lambda_g f = f$ for all $g \varepsilon G$. Thus all constant functions on V are invariant. If F is any rational function on V, the set of all $g \varepsilon G$ such that $\lambda_g F = F$ clearly forms a closed subset of G, so to prove that F is invariant it suffices to prove that $\lambda_g F = F$ for each point g in a dense subset of G. If $F \varepsilon k(V)$ and there exist generic points $g_1, \cdots, g_n$ of the various components of G over k such that $\lambda_{g_i} F = F$ for each i, then F is invariant, for in this case $\lambda_g F = F$ for any g that is generic for a component of G over k and the set of such g's is dense in G.

We now show that if K is an extension field of k and F an invariant function in $K(V)$, then F is a rational function with coefficients in K of invariant functions in $k(V)$. To do this, it suffices to assume K a finite extension of k and hence that K is either a finite algebraic extension of k or

a simple transcendental extension of k. In the former case, let $\alpha_1, \cdots, \alpha_r$ be a basis for the vector space K over k and write $F = \alpha_1 f_1 + \cdots + \alpha_r f_r$, where each $f_i \varepsilon k(V)$. If g is generic for a component of G over K we have

$$\alpha_1(f_1 - \lambda_g f_1) + \cdots + \alpha_r(f_r - \lambda_g f_r) = 0,$$

whence, by the linear disjointness over k of K and $k(g)(V)$, $f_i - \lambda_g f_i = 0$ for each i, proving that each f_i is invariant. Finally, let $K = k(t)$, where t is transcendental over k, and write

$$F = (\sum_{i \geqq 0} f_i t^i)/(\sum_{i \geqq 0} h_i t^i),$$

where each $f_i, h_i \varepsilon k(V)$, with at least one of the f_i's or h_i's a nonzero constant, and where the polynomials $\sum f_i t^i$ and $\sum h_i t^i$ are relatively prime in the ring $k(V)[t]$. Let g be a generic point over k of a component of G such that t is transcendental over $k(g)$. Then $\sum f_i t^i$ and $\sum h_i t^i$ are relatively prime in the ring $k(g)(V)[t]$. Since

$$(\sum f_i t^i)(\sum \lambda_g h_i t^i) = (\sum h_i t^i)(\sum \lambda_g f_i t^i),$$

$\sum f_i t^i$ divides $\sum \lambda_g f_i t^i$, whence there exists $c \varepsilon k(g)(V)$ such that $\sum \lambda_g f_i t^i = c \sum f_i t^i$ and $\sum \lambda_g h_i t^i = c \sum h_i t^i$. One of the f_i's or h_i's being a nonzero constant, we have $c = 1$. Hence each f_i, h_i is invariant, proving the contention of this paragraph.

We now give a method for constructing invariant functions on V. Assume first that V is a variety in a projective space. Denote by $G_1, \cdots, G_r$ the different components of G, let v be a generic point of V over k, let g_i be a generic point of G_i over $k(v)$, $(i = 1, \cdots, r)$, and let Γ_i be the locus over $k(v)$ of the point $g_i v$. Clearly Γ_i is independent of the choice of g_i. Given any $i, j = 1, \cdots, r$, we can find a point $\alpha \varepsilon G$ that is algebraic over k and such that $g_i \alpha \varepsilon G_j$; then $g_i \alpha$ is generic for G_j over $k(v)$ and Γ_j is the locus over $k(v)$ of $(g_i \alpha)v = g_i \alpha v$. Γ_i, Γ_j are also the loci of $g_i v$, $g_i(\alpha v)$ respectively over $k(\alpha, v) = k(\alpha, \alpha v)$, and since there exists a $k(\alpha)$-automorphism of the universal domain which sends v into αv it follows that Γ_j and Γ_i have the same dimension and order. Let $F_1(v), \cdots, F_N(v) \varepsilon k(v)$ be the ratios of the Chow coordinates of the cycle $\Gamma_1 + \cdots + \Gamma_r$. If g is generic for a component of G over $k(v, g_1, \cdots, g_r)$, the loci of $g_1 g^{-1} v, \cdots, g_r g^{-1} v$ over $k(v)$ are clearly $\Gamma_1, \cdots, \Gamma_r$ in some order, whence for each $\nu = 1, \cdots, N$ we have $F_\nu(g^{-1}v) = F_\nu(v)$. Considering v to be a variable point of V, we get that each F_ν is an invariant function on V. Now all points of the bunch of varieties $\Gamma_1 \cup \cdots \cup \Gamma_r$ except possibly for those lying on a certain bunch of subvarieties of smaller dimension are of the form gv, for a suitable $g \varepsilon G$

such that gv is defined (cf. [4, Appendix, Props. 8, 10]). It follows that if v_1, v_2 are generic points of V over k such that $F_\nu(v_1) = F_\nu(v_2)$ for each ν, then there exist points g_1, g_2 that are generic for certain components of G over $k(v_1, v_2)$ such that $g_1 v_1 = g_2 v_2$. If V is not imbedded in a projective space we get a similar result by replacing V by a projective model of V over k.

THEOREM 2. *Let the algebraic group G operate on the variety V and let k be a field of definition for G, V, and the operation of G on V. Then there exists a variety W and a generically surjective rational map $\tau\colon V \to W$, both W and τ being defined over k, which are characterized to within a birational correspondence defined over k by the following properties: $\tau\colon V \to W$ is generically surjective and separable, and if v_1, v_2 are generic points of V over k, then $\tau v_1 = \tau v_2$ if and only if there exist g_1, g_2 generic for components of G over $k(v_1, v_2)$ such that $g_1 v_1 = g_2 v_2$. If W' is any variety birationally equivalent to W, τ' the corresponding rational map from V to W', K any field of definition for V, W', and τ', and if we identify in the natural way functions on W' with functions on V, then $K(W')$ is the subfield of invariant functions of $K(V)$. If G is connected, then $K(W')$ is algebraically closed in $K(V)$.*

Let W be a variety defined over k such that $k(W)$ is k-isomorphic to the subfield of invariant functions of $k(V)$ and let τ be the natural rational map from V to W. Letting $\bar{k}$ denote the algebraic closure of k, the field $\bar{k}(W)$ is then the subfield of invariant functions of $\bar{k}(V)$. Since the points of G that are rational over $\bar{k}$ are dense in G, $\bar{k}(W)$ is the subfield of $\bar{k}(V)$ consisting of all functions left fixed by each automorphism λ_g of $\bar{k}(V)$, g ranging over the points of G that are rational over $\bar{k}$. By the first lemma of Section 3 of [2], $\bar{k}(V)$ is separably generated over $\bar{k}(W)$; hence τ is separable. In virtue of what has been done above, the first statement of the theorem is proved, except for the unicity part. So let W_1 and $\tau_1\colon V \to W_1$ have the requisite properties. Then each element of $k(W_1)$ is invariant, so $k(W_1) \subset k(W)$, and there exists a rational map $\sigma\colon W \to W_1$ defined over k such that $\tau_1 = \sigma\tau$. σ is one-one for generic points of W over k, so if v is generic for V over k, then $k(\tau v)$ is purely inseparable over $k(\tau_1 v)$. Since $k(v)$ is separably generated over $k(\tau_1 v)$, we must have $k(\tau v) = k(\tau_1 v)$, proving the birationality of σ and the first part of the theorem. Now let W', τ', and K be as in the theorem. Then each function of $K(W')$ is an invariant function of $K(V)$. We thus have to show that any given invariant function of $K(V)$ is a function in $K(W')$. For this, it suffices to take K finitely generated

over the prime field, so that we may form the composed field kK. The subfield of invariant functions of $kK(V) = K(k(V))$ is $K(k(W)) = kK(W')$. Hence the invariant functions of $K(V)$ are $kK(K(W')) \cap K(V)$. But kK and $K(V)$ are linearly disjoint over K, so (e.g., using [1, § 2, Lemma 3]) the latter intersection is $K(W')$. Finally, if $F \varepsilon K(V)$ is algebraic over $K(W')$ then for any $g \varepsilon G$, $\lambda_g F$ is a conjugate of F over $K(W')$, so the closed subset of G consisting of all $g \varepsilon G$ such that $\lambda_g F = F$ is a subgroup of finite index of G. If G is connected, this subgroup must be G itself, whence $F \varepsilon K(W')$. This completes the proof.

If G operates regularly on V, then the *orbit* of a point $p \varepsilon V$ is the set of all points of the form gp, g ranging over the points of G. In this case if v_1, v_2 are generic points of V over k, the condition that $\tau v_1 = \tau v_2$ is simply that v_1 and v_2 have the same orbit. Whether or not G operates regularly on V, we call the variety W, defined to within a birational transformation, the *variety of G-orbits on V*. Strictly speaking, W is a true variety of orbits only so far as *generic* orbits are concerned.

2. Algebraic subgroups and factor groups. If G is an algebraic group and H a subgroup of G, we say that H is an *algebraic subgroup* of G if H is a closed subset of G. In this case, if k is a field of definition for G and the various components of H and if h_1, h_2 are independent generic points over k of components of H passing through e, then h_1, h_2 are each specializations over k of the point $h_1h_2 \varepsilon H$. Thus H contains only one component H_0 passing through e and any other component of H is a coset of H_0. H is therefore an algebraic group. If H is a normal subgroup of G, so is H_0.

If G is an algebraic group, and $W_1, \cdots, W_r$ are subvarieties of G that pass through e, then the subgroup of G that is generated by the points of $W_1, \cdots, W_r$ is a connected algebraic subgroup of G; furthermore, this algebraic subgroup is defined over any field of definition for G and each W_i, and is complete if each W_i is complete. (For the easy proof, cf. [1, Theorem 6]). For example, any algebraic group G contains a *maximal* connected complete algebraic subgroup that contains all other connected complete algebraic subgroups of G. Similarly, if G is a connected algebraic group that is defined over k and if x, y are independent generic points of G over k then the locus W of $xyx^{-1}y^{-1}$ over k generates a connected algebraic subgroup G' of G, G' also being defined over k. Since any point of G' can be represented as the product of two generic point of G' over k, since any generic point of G' over k is a product of generic points of W over k, and since any generic point of W over k is a commutator of elements of G, we get that *the commu-*

tator subgroup of a connected algebraic group is a connected algebraic subgroup having the same field of definition.

By a *rational homomorphism* of an algebraic group G_1 into an algebraic group G_2 we mean a homomorphism of G_1 into G_2 which is given on each component of G_1 by an everywhere defined rational map of this component into some component of G_2. By a *biregular isomorphism* of the algebraic groups G_1 and G_2 we mean an isomorphism which is given by a set of biregular birational correspondences between the components of G_1 and those of G_2.

THEOREM 3. *Let G_1, G_2 be connected algebraic groups defined over k and let τ be a map of the generic points of G_1 over k into points of G_2 that satisfies the following conditions: if x and y are any independent generic points of G_1 over k then τx is rational over $k(x)$, $(y, \tau y)$ is a specialization over k of $(x, \tau x)$, and $\tau(xy) = \tau x \cdot \tau y$. Then τ can be extended in one and only one way to a homomorphism of G_1 into G_2, and the extended map τ is a rational map, defined over k, that is everywhere defined on G_1. If W is the locus of τx over k, then W is an algebraic subgroup of G_2 and τ is a surjective homomorphism from G_1 to W. The kernel H of τ is a k-closed normal algebraic subgroup of G_1 of dimension* $\dim G_1 - \dim W$.

COROLLARY. *Two birationally equivalent connected algebraic groups whose laws of composition correspond under the birational equivalence are biregularly isomorphic.*

This theorem and its corollary are the same as Theorem 5 and Corollary of [1], reproduced here because of their frequent future application. They can be extended without any difficulty to algebraic groups that are not connected. For example, if τ is a rational homomorphism of *any* algebraic group G_1 into an algebraic group G_2 then the kernel and image of τ are algebraic subgroups of G_1, G_2 respectively, and τG_1 is defined over any field of definition for G_1, G_2, and τ. Also, if the algebraic group H is embedded birationally and isomorphically in the algebraic group G, then H is a closed subset of G.

A suggestion of Weil is responsible for the following general proposition.

PROPOSITION 1. *Let V_1, V_2 be homogeneous spaces with respect to the connected algebraic group G and τ a rational map from V_1 to V_2 such that if g, v are independent generic points of G, V_1 respectively over some field of definition for G, V_1, V_2, the operation of G on V_1 and V_2, and τ, then $\tau(gv) = g(\tau v)$. Then τ is an everywhere defined surjective map from V_1 to*

V_2 and the relation $\tau(gv) = g(\tau v)$ holds for all $g \varepsilon G$, $v \varepsilon V_1$. If T is the graph of τ on $V_1 \times V_2$ and W_2 is any subvariety of V_2, then the cycle $W_1 = \mathrm{pr}_{V_1}((V_1 \times W_2) \cdot T)$ is defined, has dimension ($\dim V_1 - \dim V_2 + \dim W_2$), *the point set $|W_1|$ consists precisely of all points of V_1 that are mapped into W_2 by τ, and the map induced by τ from any component of W_1 to W_2 is surjective. If τ is separable, then each component of W_1 has coefficient one and the rational map induced by τ on each component of W_1 is also separable.*

Let k be a field of definition for G, V_1, V_2, the operation of G on V_1 and V_2, and τ. Given $p \varepsilon V_1$, choose g generic for G over $k(p)$. Then $g^{-1}p$ is generic for V_1 over k, so $\tau(g^{-1}p)$ is defined, and hence also $g(\tau(g^{-1}p))$. But when v is generic for V_1 over $k(p, g)$, we have $g(\tau(g^{-1}v)) = \tau(gg^{-1}v) = \tau(v)$, so $\tau(p) = g(\tau(g^{-1}p))$ is defined. Hence $\tau(gv) = g(\tau v)$ for all $g \varepsilon G$, $v \varepsilon V_1$. By transitivity, $\tau V_1 = V_2$. T consists of precisely all points of the form $p \times \tau p$ ($p \varepsilon V_1$), so the restriction of pr_{V_1} to T is a biregular birational correspondence between T and V_1. Since $V_1 \times V_2$ is nonsingular, to show that $(V_1 \times W_2) \cdot T$ is defined we merely have to show that the components of $(V_1 \times W_2) \cap T$ have dimension

$$\leqq (\dim V_1 \times W_2 + \dim T - \dim V_1 \times V_2) = \dim V_1 - \dim V_2 + \dim W_2.$$

But if this were not true, then for at least one point $q \varepsilon W_2$ we would have $\dim (V_1 \times q) \cap T > \dim V_1 - \dim V_2$; by transitivity, we would then have $\dim (V_1 \times q) \cap T > \dim V_1 - \dim V_2$ for *all* $q \varepsilon V_2$, which is clearly false. It follows that $(V_1 \times W_2) \cdot T$, and hence W_1, have the correct dimension, that $\tau^{-1}\{W_2\} = |W_1|$, and that $\tau |W_1| = W_2$. Now let C be any component of W_1; we wish to show that the rational map $\tau\colon C \to W_2$ is surjective. For this purpose, fix a point $p_0 \varepsilon V_1$ and consider the map $\tau_1\colon G \to V_1$ defined by $\tau_1 g = gp_0$. Letting G operate on itself by left translation, if $g_1, g_2 \varepsilon G$ we have $\tau_1(g_1g_2) = g_1g_2p_0 = g_1(\tau_1g_2)$, so our above results apply to $\tau_1\colon G \to V_1$ and also to $\tau\tau_1\colon G \to V_2$. If C_0 is a component of the point set $\tau_1^{-1}\{C\}$, then C_0 is also a component of $(\tau\tau_1)^{-1}\{W_2\}$. The restriction of $\tau\tau_1$ to C_0 must be generically surjective to W_2, for otherwise the inverse image of a certain point of W_2 would have too large a dimension. Hence if C_0' is another component of $(\tau\tau_1)^{-1}\{W_2\}$ there exist points $g \varepsilon C_0$, $g' \varepsilon C_0'$ such that $\tau\tau_1 g = \tau\tau_1 g'$; furthermore we can suppose that C_0' is the only component of $(\tau\tau_1)^{-1}\{W_2\}$ passing through g'. Setting $\gamma = g^{-1}g'$, we get $\tau\tau_1\gamma = g^{-1}\tau\tau_1 g' = g^{-1}\tau\tau_1 g = \tau\tau_1 e$. Hence

$$\tau\tau_1 C_0\gamma = C_0\tau\tau_1\gamma = C_0\tau\tau_1 e = \tau\tau_1 C_0,$$

so $C_0\gamma \subset (\tau\tau_1)^{-1}\{W_2\}$. Since $g' \varepsilon C_0\gamma$, we have $C_0\gamma = C_0'$ and $\tau\tau_1 C_0' = \tau\tau_1 C_0$. Since all components of $(\tau\tau_1)^{-1}\{W_2\}$ have the same image under $\tau\tau_1$, we get $\tau\tau_1 C_0 = W_2$. Thus $\tau C = W_2$. It remains to prove the last two statements, so suppose τ separable. To prove that all the coefficients of W_1 are one, it suffices to prove that all the coefficients of $(V_1 \times W_2) \cdot T$ are one. Let q be a simple point of W_2. Then $V_1 \times q$ can be considered as both a $V_1 \times V_2$-cycle and a $V_1 \times W_2$-cycle, and T intersects $V_1 \times q$ properly on $V_1 \times V_2$. Furthermore, the $V_1 \times V_2$-cycle $(V_1 \times W_2) \cdot T$ is also a $V_1 \times W_2$-cycle, intersecting $V_1 \times q$ properly on $V_1 \times W_2$. By [3, Theorem 18, p. 214], any component of $(V_1 \times q) \cap T$ has the same coefficient in the cycle

$$\{(V_1 \times q) \cdot \{(V_1 \times W_2) \cdot T\}_{V_1 \times V_2}\}_{V_1 \times W_2}$$

as in the cycle $\{(V_1 \times q) \cdot T\}_{V_1 \times V_2}$. Hence to show that all the coefficients of $(V_1 \times W_2) \cdot T$ are one, it suffices to prove that each coefficient of $(V_1 \times q) \cdot T$ is one. But each $g \varepsilon G$ induces a biregular birational transformation on $V_1 \times V_2$ by the law $g(v_1 \times v_2) = gv_1 \times gv_2$ and $gT = T$. Since G is transitive on V_2, it suffices to show that each coefficient of $(V_1 \times \tau v) \cdot T$ is one, where v is generic for V_1 over k, k being as above. But $v \times \tau v$ is generic over $\overline{k(\tau v)}$ for a component of $(V_1 \times \tau v) \cdot T$, so by [3, Theorem 11, p. 161] each component of $(V_1 \times \tau v) \cdot T$ has coefficient $= [k(v) : k(\tau v)]_i = 1$. Finally, if C is a component of W_1, if the field k is also a field of definition for C and W_2, and if p is a generic point of C over k, then p is a generic point over $\overline{k(\tau p)}$ for a component of $\mathrm{pr}_{V_1}((V_1 \times \tau p) \cdot T)$. Since the latter cycle has coefficients one and is rational over $k(\tau p)$, it follows that $k(p)$ is separably generated over $k(\tau p)$. Hence the restriction of τ to C is separable.

A rational homomorphism from an algebraic group G_1 to an algebraic group G_2 is called *separable* (or *purely inseparable*) if the rational map of one component of G_1 into the corresponding component of G_2 is separable (or purely inseparable); in this case the same is true for any component of G_1.

If H is an algebraic subgroup of the algebraic group G, we shall often use the same symbol H to denote the "cycle" on G consisting of the various components of H, each taken with coefficient one. Thus if G is defined over k, we say that its algebraic subgroup H is a *rational cycle over* k if the cycle H is rational over k, i.e. if the restriction of H to each component of G is rational over k; thus H is a rational cycle over k and only if it is k-closed and all its components are defined over a separable algebraic extension of k. If H is a rational cycle over k, then (since $e \varepsilon H_0$) H_0 is left fixed by all k-automorphisms of the universal domain, so H_0 is a rational cycle over k, and thus H_0 is defined over k.

COROLLARY. *If $\tau\colon G \to G'$ is a surjective separable rational homomorphism from the algebraic group G to the algebraic group G', if G, G', τ are defined over k, and if H' is an algebraic subgroup of G' that is a rational cycle over k, then the algebraic subgroup H of G which is the inverse image under τ of H' is also a rational cycle over k. In particular, the kernel of τ is a rational cycle over k.*

For if G_α is any component of G, H_α' the restriction of H' to τG_α, H_α the restriction of H to G_α, and T_α is the graph of the map $\tau\colon G_\alpha \to \tau G_\alpha$, then we can consider G_0 to operate on both G_α and τG_α, so

$$H_\alpha = \mathrm{pr}_{G_\alpha}((G_\alpha \times H_\alpha') \cdot T_\alpha)$$

is a rational cycle over k.

We digress a bit to prove a few facts needed later. First, *if the variety V is defined over k, then the points of V that are separably algebraic over k are dense in V.* This well-known result amounts to showing that if $(x_1,\cdots,x_n)$ are quantities such that $k(x_1,\cdots,x_n)$ is separably generated over k and $F(X) \varepsilon K[X_1,\cdots,X_n]$ (K some extension field of k), $F(x) \neq 0$, then there exists a finite k-specialization $(\bar{x}_1,\cdots,\bar{x}_n)$ of $(x_1,\cdots,x_n)$ such that $F(\bar{x}) \neq 0$ and such that $k(\bar{x})$ is separably algebraic over k. Assuming $x_1,\cdots,x_r$ to be a separating transcendence basis of $k(x)$ over k and replacing k by $k(x_1,\cdots,x_{r-1})$ we reduce this to the case where $k(x)$ has transcendence degree 1 over k, in which case it is known that there exists an infinity of places of $k(x)$ over k whose residue fields are separable over k. (Our contention can also be deduced immediately from [4, Appendix, Prop. 13]). We now claim that if we have the situation described in Prop. 1, with τ separable, and if k is a field of definition for G, V_1, V_2, the operation of G on V_1 and V_2, and τ, then *for any point $p_2 \varepsilon V_2$ there exists a point $p_1 \varepsilon V_1$ such that $\tau p_1 = p_2$ and $k(p_1)$ is separably algebraic over $k(p_2)$.* For each component of the cycle $\mathrm{pr}_{V_1}((V_1 \times p_2)\cdot T)$ is defined over a separable algebraic extension of $k(p_2)$, so we may use the preceding remark.

LEMMA. *Let the algebraic group G operate on the variety V and let k be a field of definition for G, V, and the operation of G on V. Then any algebraic subgroup H of G operates on V, and if H is a rational cycle over k then the variety W of H-orbits on V and the natural rational map $\tau\colon V \to W$ may both be taken to be defined over k.*

By the lemma to Theorem 1, H operates on V. So let H as a cycle be rational over k, and let the overfield k' of k be a finite separable normal

algebraic extension of k over which each component of H is defined. Let $\alpha_1, \cdots, \alpha_n$ be a basis for k' over k and let $\sigma_1, \cdots, \sigma_n$ be the distinct k-automorphisms of k'. Then any H-invariant function in $k'(V)$ is of the form $\sum \alpha_i F_i$, where each $F_i \varepsilon k(V)$. Since H is a rational cycle over k, for any $j = 1, \cdots, n$, $\sum \sigma_j(\alpha_i) F_i$ is also an H-invariant function in $k'(V)$. Since $|\sigma_j(\alpha_i)| \neq 0$, each F_i is H-invariant. By Theorem 2, the lemma is proved by taking W to be any variety defined over k such that $k(W)$ is k-isomorphic to the subfield of H-invariant functions of $k(V)$, and by letting τ be the natural rational map from V to W.

THEOREM 4. *Let G be an algebraic group and H a normal algebraic subgroup of G. Then there exists an algebraic group G/H and a surjective separable rational homomorphism τ: $G \to G/H$ with kernel H. These properties characterize G/H and τ to within a biregular isomorphism. Also, if k is a field of definition of G over which the cycle H is rational, then G/H and τ may be taken to be defined over k.*

This is proved by Weil [5, Prop. 2] for the case when G is connected. His proof can be extended to cover the general case, but for the sake of completeness we give a direct proof by a silghtly different method.

Let G, H, k be as above. We first prove the existence of G/H and τ in the special case where $H \subset G_0$. For any component G_α of G, G_0 operates regularly on G_α by the rule $g_0 \times g_\alpha \to g_\alpha g_0^{-1}$. Consider the induced operation of H on G_α, and construct the variety W_α of H-orbits on G_α and the natural rational map τ_α from G_α to W_α, W_α and τ_α being taken (by the lemma) to be defined over k. We wish to show first that the other obvious operation of G_0 on G_α, namely $g_0 \times g_\alpha \to g_0 g_\alpha$, induces an operation of G_0 on W_α. Let x_0 and x_α be independent generic points over k of G_0 and G_α respectively, and let $f(x_0, x_\alpha) \varepsilon k(\tau_\alpha(x_0 x_\alpha))$. Then if h is generic for a component of H over $\bar{k}(x_0, x_\alpha)$ we have $\tau_\alpha(x_0 x_\alpha h^{-1}) = \tau_\alpha(x_0 x_\alpha)$, so $f(x_0, x_\alpha h^{-1}) = f(x_0, x_\alpha)$. Imagining x_0 fixed and x_α variable on G_α, $f(x_0, x_\alpha)$ becomes an H-invariant function on G_α that is defined over $k(x_0)$, whence $f(x_0, x_\alpha) \varepsilon k(x_0, \tau_\alpha x_\alpha)$. Thus $\tau_\alpha(x_0 x_\alpha)$ is rational over $k(x_0, \tau_\alpha x_\alpha)$, so we write $x_0(\tau_\alpha x_\alpha) = \tau_\alpha(x_0 x_\alpha)$. If $x_0 \times y_0 \times x_\alpha$ is generic for $G_0 \times G_0 \times G_\alpha$ over k, then

$$x_0(y_0(\tau_\alpha x_\alpha)) = x_0(\tau_\alpha(y_0 x_\alpha)) = \tau_\alpha(x_0 y_0 x_\alpha) = x_0 y_0(\tau_\alpha x_\alpha).$$

Also, since $x_0^{-1} \times x_0 x_\alpha$ is generic for $G_0 \times G_\alpha$ over k, $\tau_\alpha x_\alpha = x_0^{-1}(\tau_\alpha(x_0 x_\alpha))$ is rational over $k(x_0, x_0(\tau_\alpha x_\alpha))$. Thus G_0 operates on W_α by the law $x_0(\tau_\alpha x_\alpha) = \tau_\alpha(x_0 x_\alpha)$, and this operation is defined over k. W_α is a prehomogeneous space with respect to G_0, since $\tau_\alpha(x_0 x_\alpha)$ is generic for W_α over $k(\tau_\alpha x_\alpha)$, so by

the main result of [4] we may, without any loss of generality, assume that W_α is a homogeneous space with respect to G_0. Then Proposition 1 shows that τ_α is defined everywhere on G_α and $\tau_\alpha G_\alpha = W_\alpha$. If $p_1, p_2 \varepsilon G_\alpha$ and x_0 is generic for G_0 over $k(p_1, p_2)$, we have $\tau_\alpha p_1 = \tau_\alpha p_2$ if and only if $\tau_\alpha(x_0 p_1) = \tau_\alpha(x_0 p_2)$, which is true if and only if $x_0 p_1$ and $x_0 p_2$ have the same H-orbit on G_α, i. e. $p_1^{-1} p_2 \varepsilon H$. Thus τ_α is a one-one map from left cosets of H on G_α to W_α. (Note that in the case $G = G_0$, we have constructed the homogeneous left coset space G/H. The normality of H has not yet been used.) Now let G_α, G_β be components of G, let x_α, x_β be independent generic points over k of G_α, G_β respectively, suppose $x_\alpha x_\beta^{-1} \varepsilon G_\gamma$, and let $f(x_\alpha, x_\beta) \varepsilon k(\tau_\gamma(x_\alpha x_\beta^{-1}))$. Imagining x_α fixed and x_β variable on G_β, f becomes an H-invariant function in $k(x_\alpha)(G_\beta)$, hence in $k(x_\alpha)(W_\beta)$; i. e. $f \varepsilon k(x_\alpha, \tau_\beta x_\beta)$. Similarly, imagining x_β fixed and x_α variable on G_α, f becomes an H-invariant function in $k(\tau_\beta x_\beta)(G_\alpha)$, and hence in $k(\tau_\beta x_\beta)(W_\alpha)$. Thus $f \varepsilon k(\tau_\alpha x_\alpha, \tau_\beta x_\beta)$. Therefore $\tau_\gamma(x_\alpha x_\beta^{-1})$ is rational over $k(\tau_\alpha x_\alpha, \tau_\beta x_\beta)$; that is, we have a rational map $\phi_{\alpha,\beta}: W_\alpha \times W_\beta \rightarrow W_\gamma$, defined by $\phi_{\alpha,\beta}(\tau_\alpha x_\alpha, \tau_\beta x_\beta) = \tau_\gamma(x_\alpha x_\beta^{-1})$. Let $g_\alpha \varepsilon G_\alpha$ and let x_0 be a generic point of G_0 over $k(g_\alpha, x_\alpha, x_\beta)$. Then the relation

$$\phi_{\alpha,\beta}(\tau_\alpha x_\alpha, \tau_\beta x_\beta) = x_0^{-1}\phi_{\alpha,\beta}(x_0(\tau_\alpha x_\alpha), \tau_\beta x_\beta)$$

implies

$$\phi_{\alpha,\beta}(\tau_\alpha g_\alpha, \tau_\beta x_\beta) = x_0^{-1}\phi_{\alpha,\beta}(x_0(\tau_\alpha g_\alpha), \tau_\beta x_\beta),$$

so $\phi_{\alpha,\beta}$ is defined whenever its second argument is generic for W_β over k. Thus if $g_\alpha \varepsilon G_\alpha$, $g_\beta \varepsilon G_\beta$, and x_0 is generic for G_0 over $k(g_\alpha, g_\beta)$, the relation

$$\phi_{\alpha,\beta}(\tau_\alpha g_\alpha, \tau_\beta g_\beta) = \phi_{\alpha,\beta}(\phi_{\alpha,0}(\tau_\alpha g_\alpha, \tau_0 x_0), \phi_{\beta,0}(\tau_\beta g_\beta, \tau_0 x_0))$$

shows that $\phi_{\alpha,\beta}$ is defined everywhere on $W_\alpha \times W_\beta$. This shows that the union $\{W_\alpha\}$ of all the W_α's (which is an abstract group in the obvious manner) is actually an algebraic group defined over k, and that the map τ, defined as τ_α on G_α, is a surjective rational homomorphism from G to $\{W_\alpha\}$ having the requisite properties. Now drop the condition $H \subset G_0$. Then $G_0 \cap H$ is a normal algebraic subgroup of G that is a rational cycle over k, so that if we first factor out $G_0 \cap H$, we are reduced to proving the existence of G/H and τ in the special case where $G_0 \cap H = \{e\}$. Here H consists of a finite number of points, each rational over k, so we may take G/H to be a set of disjoint replicas of certain components of G, exactly one component being taken from each coset of $G_0 H$, with the obvious group law on the set G/H and the obvious map τ from G to G/H. This ends the proof of the existence of G/H and τ.

For the unicity, note that if τ_1, τ_2 are surjective separable rational homo-

morphisms with kernel H from G to the algebraic groups G_1, G_2 respectively, then the unicity part of Theorem 2 (applied to the operation of $G_0 \cap H$ on G_0) shows that $(G_1)_0$ and $(G_2)_0$ are birationally equivalent. Since their group laws correspond under their birational equivalence, $(G_1)_0$ and $(G_2)_0$ are biregularly isomorphic. Hence G_1 and G_2, which are naturally isomorphic as abstract groups, are biregularly isomorphic.

Corollary 1. *G/H and τ are characterized to within a biregular isomorphism by the following properties:*

(1) *τ is a surjective rational homomorphism from G to G/H with kernel H.*

(2) *if τ' is a rational homomorphism of G into an algebraic group G' and the kernel of τ' contains H, then there exists a rational homomorphism σ of G/H into G' such that $\tau' = \sigma\tau$.*

Property (1) is known. For property (2), there exists a well-defined homomorphism $\sigma\colon G/H \to G'$ such that $\tau' = \sigma\tau$ and we need only prove σ rational. Hence we may take G connected and $G' = \tau' G$. If k is a field of definition for G, H, G/H, τ, G', τ', then $k(G')$ consists of functions in $k(G)$ that are H-invariant, so $k(G') \subset k(G/H)$. Hence σ is rational. Conversely, properties (1) and (2) clearly characterize G/H and τ to within a biregular isomorphism.

Examples. If G is an algebraic group then $G/\{e\}$ and G/G are biregularly isomorphic to G and $\{e\}$ respectively. If G and H are algebraic groups then the abstract group $G \times H$ is made into an algebraic group by identifying it with the union of the various products of components of G by components of H, and we have $(G \times H)/H$ biregularly isomorphic to G.

If H is any algebraic subgroup (not necessarily normal) of the connected algebraic group G, G/H will denote the G-homogeneous space of left cosets of H on G.

Corollary 2. *If $H \supset N$ are algebraic subgroups of the algebraic group G with N normal in G, then the natural isomorphism from H/N into G/N is biregular.*

Corollary 3. *If G is a connected algebraic group, H a connected algebraic subgroup of G, and N an algebraic subgroup of H, then the natural map from H/N into G/N is a biregular birational correspondence.*

Corollary 2 comes from Corollary 3 by restricting one's attention to the

components of the identity of G and H, so it is necessary to prove only the latter. Let $\tau\colon G \to G/N$ be the natural map. The composite of the injection of H into G and τ gives a natural everywhere defined rational map from H to τH. Letting N operate on H by the law $n(h) = hn^{-1}$, any rational function on τH induces an N-invariant function on H, so we have a natural rational map from H/N to τH. Considering G and G/N to be G-homogeneous spaces (with respect to left translation by elements of G), since τ is separable Proposition 1 shows that our map from H to τH is separable. Hence the map from H/N to τH is separable. Since the latter map is generically one-one, it is birational. Since H/N and τH are both H-homogeneous spaces (with respect to left translation by elements of H) this map from H/N to τH is biregular.

THEOREM 5. *Let G be an algebraic group that operates on the variety V and let H be a normal algebraic subgroup of G. Then G/H operates on the variety W of H-orbits on V by the rule $gH(Hv) = Hgv$ and the variety of (G/H)-orbits on W is naturally birationally equivalent to the variety of G-orbits on V.*

Let $\theta\colon G \to G/H$, $\tau\colon V \to W$ be the natural rational maps and let k be a field of definition for G, V, the operation of G on V, G/H, W, θ, and τ. Let g be generic for a component of G over k, v generic for V over $k(g)$, h generic for a component of H over $\bar{k}(g,v)$. Then $\tau(g(hv)) = \tau(ghg^{-1}(gv)) = \tau(gv)$. If we imagine g fixed and v variable we get that $\tau(gv)$ is H-invariant, so $\tau(gv)$ is rational over $k(g,\tau v)$. Also, $\tau((gh)v) = \tau(g(hv)) = \tau(gv)$, so if we imagine v fixed, g variable for a component G_α of G and h generic over $\bar{k}(v)$ for a component of $G_0 \cap H$, and if we note that θG_α is the variety of $(G_0 \cap H)$-orbits on G_α, we get that $\tau(gv)$ is rational over $k(\theta g, \tau v)$. Thus we have a generically surjective rational map from each variety $\theta G_\alpha \times W$ to W. If γ_1, γ_2 are independent generic points over k of components of G/H and w is generic for W over $k(\gamma_1, \gamma_2)$, we clearly have $\gamma_1(\gamma_2 w) = (\gamma_1\gamma_2)w$. Since $\tau v = \tau(g^{-1}(gv))$ is rational over $k(\theta g, \tau(gv))$, we get $k(\theta g, \tau(gv)) = k(\theta g, \tau v)$, so G/H operates on W. The last statement of the theorem merely says that the field of (G/H)-invariant functions on W is the field of G-invariant functions on V.

COROLLARY. *If k is a field of definition for G, V, the operation of G on V, G/H, W, the variety of (G/H)-orbits on W, the variety of G-orbits on V, and for the rational maps of G, V, W, and V respectively on the four preceding varieties, then k is a field of definition for the operation of G/H*

on W and the birational equivalence of the variety of (G/H)-orbits on W with the variety of G-orbits on V.

3. Isogeny and the homomorphism theorems. We say that the connected algebraic groups G_1 and G_2 are *isogenous* if there exists a connected algebraic group G_3 and surjective rational homomorphisms with finite kernel from G_3 to G_1 and G_2; if these rational homomorphisms are both separable, or both purely inseparable, we say that G_1 and G_2 are *separably isogenous* or *inseparably isogenous* respectively. Clearly isogenous algebraic groups have the same dimension.[2]

If one of two isogenous connected algebraic groups is commutative, then so is the other: in one direction this is trivial, in the other we use the connectedness of the commutator subgroup of a connected algebraic group.

THEOREM 6. *Isogeny is an equivalence relation among connected algebraic groups. If G and H are isogenous connected algebraic groups then there is a one-one correspondence between the connected algebraic subgroups of G and those of H such that corresponding subgroups are isogenous, and if the connected algebraic subgroups G_1 and G_2 of G correspond to the subgroups H_1 and H_2 of H, then $G_1 \supset G_2$ if and only if $H_1 \supset H_2$; if $G_1 \supset G_2$ then G_2 is normal in G_1 if and only if H_2 is normal in H_1, and in this case, G_1/G_2 and H_1/H_2 are isogenous. If G, G', H, H' are connected algebraic groups with G isogenous to H and G' isogenous to H', then $G \times G'$ is isogenous to $H \times H'$. Also, these same results hold if the notion "isogeny" is replaced throughout by either "separable isogeny" or "inseparable isogeny."*

We carry through the proof simultaneously for isogeny, separable isogeny, and inseparable isogeny. k will always denote a field of definition for all the algebraic groups and rational homomorphisms in question at any time. Obviously the relations isogeny, etc., are each reflexive and symmetric. To prove transitivity, let σ_1, σ_2 be surjective rational homomorphisms with finite

[2] In [6] two abelian varieties A and B are said to be isogenous if they have the same dimension and if there exists a surjective rational homomorphism from A to B. This is an equivalence relation for abelian varieties but not for more general groups, which accounts for the necessity of our present definition (clearly an extension of the older one). For an example in which the old definition is inadequate, let G_1 be the algebraic group of all $n \times n$ matrices of determinant 1, where $n > 1$ is not a power of the field characteristic, and let $G_2 = G_1/C$, C being the center of G_1. C consists of all multiples of the unit matrix by n-th roots of unity, hence has finite order > 1, while G_2 is the projective group. Since G_2 has no proper normal subgroups it possesses no surjective homomorphism to G_1.

kernel from the connected algebraic group G to the algebraic groups G_1, G_2 respectively, and let τ_2, τ_3 be surjective rational homomorphisms with finite kernel from the connected algebraic group G' to the algebraic groups G_2, G_3 respectively. Let $\sigma_2 \times \tau_2$ be the obvious surjective homomorphism from $G \times G'$ to $G_2 \times G_2$ and let Γ be the component of the identity of the inverse image under $\sigma_2 \times \tau_2$ of the diagonal on $G_2 \times G_2$. Since $\sigma_2 \times \tau_2$ has finite kernel, Γ has the same dimension as G_2. If $p \times p'$ is a generic point of Γ over k, then p, p' are generic over k for G, G' respectively, $\sigma_2 p = \tau_2 p'$, and $k(p \times p')$ is algebraic over $k(\sigma_2 p)$. The maps $p \times p' \rightarrow p$, $p \times p' \rightarrow p'$ define surjective rational homomorphisms with finite kernel from Γ to G, G' respectively, and hence we have surjective rational homomorphisms with finite kernel from Γ to both G_1 and G_3. In the case of separable (or inseparable) isogeny, $k(p \times p')$ is separable (or purely inseparable) over $k(\sigma_2 p)$ and hence $k(p \times p')$ is separable (or purely inseparable) over each of the fields $k(p)$ and $k(p')$. Thus the relations isogeny, separable isogeny, and inseparable isogeny are equivalence relations. Next let σ, τ be surjective rational homomorphisms with finite kernel from the connected algebraic group Γ to the algebraic groups G, H respectively. We say that the connected algebraic subgroups G_1, H_1 of G, H respectively correspond if there exists a connected algebraic subgroup Γ_1 of Γ such that $\sigma\Gamma_1 = G_1$ and $\tau\Gamma_1 = H_1$. Since for any connected algebraic subgroup G_1 of G there exists one and only one connected algebraic subgroup Γ_1 of Γ such that $\sigma\Gamma_1 = G_1$ (namely $\Gamma_1 =$ component of the identity of $\sigma^{-1}(G_1)$), we get the one-one correspondence claimed by the theorem, and the corresponding groups G_1, H_1 are clearly isogenous. In the case of separable isogeny we can apply Proposition 1 to the separable map $\sigma \colon \Gamma \rightarrow G$; this shows that the rational homomorphism from Γ_1 to G_1 is separable, so corresponding subgroups G_1, H_1 are separable isogenous. In the case of inseparable isogeny, σ and τ are isomorphisms, so the homomorphisms from Γ_1 to G_1 and H_1 are isomorphisms, hence purely inseparable. If G_1, G_2 correspond to H_1, H_2, then clearly $G_1 \supset G_2$ if and only if $H_1 \supset H_2$. For the statements about normality and factor groups, it suffices to take $G_1 = G$, $H_1 = H$. Then if G_2 is normal in G, $\sigma^{-1}(G_2)$ is normal in Γ, so the component of the identity of $\sigma^{-1}(G_2)$ is normal in Γ, whence H_2 is normal in H. If Γ_1 is a connected normal algebraic subgroup of Γ and $G_1 = \sigma\Gamma_1$, $H_1 = \tau\Gamma_1$, then we have surjective rational homomorphisms from Γ to G/G_1 and H/H_1. The kernels of these homomorphisms contain Γ_1, so we get surjective rational homomorphisms from Γ/Γ_1 to G/G_1 and H/H_1. The kernels of these last two homomorphisms being finite, G/G_1 and H/H_1 are isogenous. In the case of separable isogeny, all the above homomorphisms

are separable, so G/G_1 and H/H_1 are separably isogenous. In the case of inseparable isogeny, the homomorphisms from Γ/Γ_1 to G/G_1 and H/H_1 are isomorphisms, so G/G_1 and H/H_1 are inseparably isogenous. The statement about direct products is immediate.

COROLLARY 1. *If G and H are isogenous connected algebraic groups and if $G = G_0 \supset G_1 \supset G_2 \supset \cdots$ is a normal chain of connected algebraic subgroups of G, then there exists a normal chain of connected algebraic subgroups of H, say $H = H_0 \supset H_1 \supset H_2 \supset \cdots$ such that each G_i is isogenous to H_i and each G_i/G_{i+1} is isogenous to H_i/H_{i+1}. The same result holds if we replace "isogeny" by either "separable isogeny" or "inseparable isogeny."*

If G_1 and G_2 are algebraic groups (not necessarily connected), we say that G_1 and G_2 are *inseparably isogenous* if there exists an algebraic group G_3 and surjective rational isomorphisms from G_3 to G_1 and G_2. This agrees with the previous definition in the case of connected algebraic groups. The following result is got by trivially modifying the above proofs.

COROLLARY 2. *Theorem 6 and Corollary 1 hold for inseparable isogeny if we delete the word "connected."*

In characteristic zero inseparable isogeny is the same as biregular isomorphism. In characteristic $p \neq 0$, consider the automorphism Θ of the universal domain defined by $\Theta\alpha = \alpha^p$. Under this automorphism a point in affine space with coordinates $(\alpha_1, \cdots, \alpha_r)$ will go into the point $(\alpha_1{}^p, \cdots, \alpha_r{}^p)$, a variety V will go into a variety ΘV, and an algebraic group G will go into an inseparably isogenous algebraic group ΘG. If τ is a surjective rational isomorphism from the algebraic group G to the algebraic group G', then for n sufficiently large the isomorphism $\Theta^n\tau^{-1}$ from G' to $\Theta^n G$ will be rational. Hence if G_1, G_2 are inseparably isogenous algebraic groups the obvious surjective isomorphism from G_1 to $\Theta^n G_2$ will be rational for n sufficiently large.

THEOREM 7. *If $H \supset N$ are algebraic subgroups of the connected algebraic group G with N normal in G, then the natural one-one correspondence between the points of G/H and $(G/N)/(H/N)$ is birational and biregular. If G is not necessarily connected but both H and N are normal in G, then the natural isomorphism between G/H and $(G/N)/(H/N)$ is biregular.*

The second part of the theorem follows from the first by restricting one's attention to G_0, $G_0 \cap H$ and $G_0 \cap N$. To prove the first part, note first that we have a sequence of natural separable rational maps

$$G \to G/N \to (G/N)/(H/N),$$

so the natural map from G to $(G/N)/(H/N)$ is a separable rational map. Bearing in mind that G/H is the variety of H-orbits on G (H operating on G by the law $h(g) = gh^{-1}$), and similarly for G/N, H/N, and $(G/N)/(H/N)$, we see that any rational function on $(G/N)/(H/N)$ gives rise to an H/N-invariant function on G/N, and thence to an H-invariant function on G. Thus there exists a natural generically surjective rational map from G/H to $(G/N)/(H/N)$. This last map being separable and generically one-one, it is actually birational. Since both G/H and $(G/N)/(H/N)$ are G-homogeneous spaces (G operating by left translation), the birational correspondence between G/H and $(G/N)/(H/N)$ is biregular.

THEOREM 8. *Let H and N be algebraic subgroups of the algebraic group G, with N normal in G. Then HN is an algebraic subgroup of G and the natural surjective isomorphism from $H/(H \cap N)$ to HN/N is rational.*

By the comments at the beginning of Section 2, the subgroup H_0N_0 of G is algebraic; hence HN is an algebraic subgroup of G. The natural isomorphism from H into HN is rational, so the homomorphism from H to HN/N is rational. The kernel of this homomorphism being $H \cap N$, we have a rational isomorphism from $H/(H \cap N)$ to HN/N.

If V is a variety and $\tau\colon V \rightarrow W$ a generically surjective rational map, k a field of definition for V, W, and τ, and x a generic point of V over k, the *order of inseparability of* τ is $[k(x):k(\tau x)]_i$; this is independent of the choice of x and, by [3, Theorem 11, p. 161], independent of the field k. If τ is a surjective rational homomorphism from an algebraic group G to an algebraic group G', by the *order of inseparability of* τ we mean the order of inseparability of the restriction of τ to any component of G; this is independent of the component chosen.

The following proposition, which gives more precise information on Theorem 8, will not be used in the sequel.

PROPOSITION 2. *Let H, N be algebraic subgroups of the connected algebraic group G, with N normal in G and $HN = G$. Then H and N intersect properly on G and $H \cdot N = q(H \cap N)$, where q is the order of inseparability of the natural rational isomorphism from $H/(H \cap N)$ to HN/N. If k is a field of definition for H, N and G, and h and n are independent generic points of components of H and N respectively over k, then $q = [k(h,n):k(hn)]_i$.*

Let τ_1 be the homomorphism from H to $H/(H \cap N)$ and τ the homomorphism from G to $HN/N = G/N$. Then if k, h, n are as above, the

order of inseparability q of the homomorphism from $H/(H \cap N)$ to HN/N is given by $q = [k(\tau_1 h) : k(\tau h)]_i$. Since τ_1 is separable, [3, Prop. 27, p. 23] gives $q = [k(h) : k(\tau h)]_i$. Using the paragraph after Prop. 1, Cor., choose $\bar{h} \varepsilon G$ such that $\tau\bar{h} = \tau h$ and $k(\bar{h})$ is a separable algebraic extension of $k(\tau h)$. By [3, Prop. 25, p. 22], $q = [k(h, \bar{h}) : k(\bar{h})]_i$. The point $\bar{h}^{-1}hn$ is generic for a component of N over $k(h, \bar{h})$, $k(\bar{h}, \bar{h}^{-1}hn)$ is a regular extension of $k(\bar{h})$, and the fields $k(h, \bar{h})$ and $k(\bar{h}, \bar{h}^{-1}hn)$ are free with respect to each other over $k(\bar{h})$, hence linearly disjoint over $k(\bar{h})$. By [3, Prop. 26, p. 23], $q = [k(h, \bar{h}, \bar{h}^{-1}hn) : k(\bar{h}, \bar{h}^{-1}hn)]_i = [k(h, n, \bar{h}) : k(hn, \bar{h})]_i$. Since $k(\bar{h})$ is separably algebraic over $k(\tau h) = k(\tau(hn))$, [3, Proposition 25, p. 22] gives $q = [k(h, n) : k(hn)]_i$. Next, the equality of the dimensions of $H/(H \cap N)$ and G/N implies that the cycles H and N intersect properly on G. Since there exists a translation on G taking any given component of $H \cap N$ into any other given component, these components all occur to the same multiplicity in $H \cdot N$. We shall complete the proof by showing that this common multipilicity is $[k(h, n) : k(hn)]_i$. But we could have taken h, n to lie in H_0, N_0 respectively, so from now on we may assume that H and N are *connected.* If $p \varepsilon G$ we can write $p = h_1 n_1$, with $h_1 \varepsilon H$, $n_1 \varepsilon N$, and therefore $H \cap pN = H \cap h_1 N = h_1(H \cap N)$. Thus H and pN intersect properly on G and the common multiplicity of the components of $H \cap N$ in $H \cdot N$ equals the multiplicity of any component of $H \cap pN$ in $H \cdot pN$. But by [6, Cor. 1, p. 24], $p \times H \cdot pN = W \cdot (p \times G)$, where $W \subset G \times G$ is the locus over k of $hn^{-1} \times h$. W is also the locus over k of $hn \times h$. Taking p to be the point hn and applying [3, Theorem 11, p. 161], we get that each component of $W \cdot (p \times G)$ has multiplicity $[k(hn, h) : k(hn)]_i = [k(h, n) : k(hn)]_i$. This completes the proof.

COROLLARY. *Let H, N be connected algebraic subgroups of the algebraic group G and suppose that almost all points of G are of the form hn, with $h \varepsilon H$ and $n \varepsilon N$. Then the map $\tau \colon H \times N \to G$ defined by $\tau(h \times n) = hn$ is birational if and only if $H \cdot N = e$. If N is a normal subgroup of G and $H \cdot N = e$, then τ is biregular.*

Let k be a field of definition for H, N, and G and let $h \times n$ be a generic point of $H \times N$ over k. Then hn is generic for G over k. Note that the last part of the proof of the proposition made no use of the normality of N, and in fact showed that if H and N intersect properly on G, then $H \cdot N = [k(h, n) : k(hn)]_i (H \cap N)$. If τ is birational, it must be one-one almost everywhere, so $H \cap N = \{e\}$. Furthermore, in this case $\dim G = \dim H + \dim N$, so $H \cdot N$ is defined. Since $k(hn) = k(h, n)$, we have

$H \cdot N = e$. Conversely, let $H \cdot N = e$. Since τ is one-one almost everywhere, $k(h,n)$ is purely inseparable over $k(hn)$. Since $[k(h,n):k(hn)]_i = 1$, we have $k(h,n) = k(hn)$, so τ is birational. If in addition N is normal in G then τ is one-one and surjective, so Zariski's main theorem on birational correspondences gives the biregularity of τ.

We now prove the analogues for algebraic groups of the Zassenhaus lemma and the Jordan-Hölder-Schreier theorem.

LEMMA. *Let $H_1 \supset N_1$ and $H_2 \supset N_2$ be algebraic subgroups of the algebraic group G, with N_1, N_2 normal in H_1, H_2 respectively. Then $N_1(H_1 \cap N_2)$ and $N_2(H_2 \cap N_1)$ are normal algebraic subgroups of the algebraic groups $N_1(H_1 \cap H_2)$ and $N_2(H_2 \cap H_1)$ respectively and the algebraic groups*

$$\frac{N_1(H_1 \cap H_2)}{N_1(H_1 \cap N_2)} \quad \text{and} \quad \frac{N_2(H_2 \cap H_1)}{N_2(H_2 \cap N_1)}$$

are naturally isomorphic, the isomorphism being an inseparable isogeny.

By Theorem 8, $N_1(H_1 \cap N_2)$ and $N_1(H_1 \cap H_2)$ are algebraic subgroups of H_1, and the first is normal in the second. There exists a natural rational isomorphism from $H_1 \cap H_2$ into $N_1(H_1 \cap H_2)$ and hence there exists a natural surjective rational homomorphism from $H_1 \cap H_2$ to

$$N_1(H_1 \cap H_2)/N_1(H_1 \cap N_2).$$

But the kernel of this last homomorphism is

$$(H_1 \cap H_2) \cap (N_1(H_1 \cap N_2)) = (H_1 \cap N_2)(H_2 \cap N_1).$$

Hence there exists a natural surjective rational isomorphism from

$$(H_1 \cap H_2)/(H_1 \cap N_2)(H_2 \cap N_1) \text{ to } N_1(H_1 \cap H_2)/N_1(H_1 \cap N_2).$$

The whole lemma now follows from symmetry.

THEOREM 9. *If G is an algebraic group, then any two normal chains of algebraic subgroups of G have refinements consisting of normal chains of algebraic subgroups whose successive algebraic factor groups, except for order, are inseparably isogenous.*

For if

$$G = G_0 \supset G_1 \supset \cdots \supset G_r = \{e\} \text{ and } G = H_0 \supset H_1 \supset \cdots \supset H_s = \{e\}$$

are the given normal chains, the standard proof of the algebraic version of this theorem (cf. Zassenhaus, *Lehrbuch der Gruppentheorie*) consists in inter-

posing in the first chain groups of the form $G_i(G_{i-1} \cap H_j)$ and analogous groups in the second chain, so it all reduces to the lemma.

If $G = G_0 \supset G_1 \supset \cdots \supset G_r = \{e\}$ is a normal chain of algebraic subgroups of the algebraic group G, and if H_i is the component of the identity of G_i $(i = 0, \cdots, r)$, then $H_0 \supset H_1 \supset \cdots \supset H_r = \{e\}$ is a normal chain of algebraic subgroups of H_0 and each H_{i-1}/H_i $(i = 1, \cdots, r)$ is separably isogenous to the component of the identity of G_{i-1}/G_i. Define an algebraic group to be *simple* if it contains no proper connected normal algebraic subgroup. The following results immediately from Theorem 9 and Theorem 6, Cor. 1.

COROLLARY. *If G and H are isogenous connected algebraic groups and if $G = G_0 \supset G_1 \supset \cdots \supset G_r = \{e\}$ and $H = H_0 \supset H_1 \supset \cdots \supset H_s = \{e\}$ are normal chains of connected algebraic subgroups such that each G_{i-1}/G_i and each H_{i-1}/H_i is simple and of dimension > 0, then $r = s$ and the successive algebraic factor groups of G and H are isogenous, except for order.*

4. Solvable algebraic groups and cross sections. Throughout the remainder of this paper G_a will denote the algebraic group consisting of the affine line with the composition law $(x_1)(x_2) = (x_1 + x_2)$ and G_m will denote the algebraic group of the affine line minus the origin with the composition law $(x_1)(x_2) = (x_1 x_2)$. Any connected noncomplete algebraic group of dimension one is biregularly isomorphic to either G_a or G_m. If p is the field characteristic, then G_m has elements of any given finite order not divisible by p and no element of order p, while each element $\neq e$ of G_a has order p if $p \neq 0$ and e is the only element of G_a having finite order if $p = 0$. A rational image of a rational curve being rational, any rational homomorphic image with finite kernel of G_a or G_m is biregularly isomorphic to G_a or G_m respectively. Using the fact that if an algebraic group G has a noncomplete rational homomorphic image then G itself is noncomplete, we get that any connected algebraic group that is isogenous to G_a or G_m is biregularly isomorphic to G_a or G_m respectively, while G_a and G_m are themselves nonisogenous.

We say that an algebraic group is *solvable* if it has a normal chain of algebraic subgroups such that each algebraic factor group is biregularly isomorphic to either G_a, G_m, or a finite commutative group. If one of two inseparably isogenous algebraic groups is solvable, so is the other. A connected solvable algebraic group has a normal chain of connected algebraic subgroups such that each algebraic factor group is biregularly isomorphic to G_a or G_m, hence if one of two isogenous connected algebraic groups is solvable,

so is the other. By standard arguments, any algebraic subgroup of a solvable algebraic group is solvable, any rational homomorphic image of a solvable algebraic group is solvable, and if an algebraic group G contains a normal solvable algebraic subgroup H such that G/H is solvable, then G is solvable.

An *algebraic group of matrices* is an algebraic subgroup of the multiplicative group of all invertible square matrices (of some degree $n > 0$) with coefficients in the universal domain. For example, the full triangular group T of degree n consisting of all invertible $n \times n$ matrices (a_{ij}) such that $a_{ij} = 0$ if $i > j$, and its subgroups T_ν $(\nu = 1, \cdots, n)$ defined by $a_{ij} = 0$ if $i > j$ or if $1 \leqq j - i < \nu$, while $a_{11} = \cdots = a_{nn} = 1$, are all connected algebraic groups of matrices, and $T \supset T_1 \supset \cdots \supset T_n = \{e\}$, where e is the unit matrix. The map $(a_{ij}) \rightarrow (a_{11}, \cdots, a_{nn})$ clearly is a surjective separable rational homomorphism from T to $(G_m)^n$ with kernel T_1, so T_1 is normal in T and $T/T_1 = (G_m)^n$. Similarly, if $1 \leqq \nu < n$, the map $(a_{ij}) \rightarrow (a_{1,\nu+1}, a_{2,\nu+2}, \cdots, a_{n-\nu,n})$ is a surjective separable rational homomorphism from T_ν to $(G_a)^{n-\nu}$ with kernel $T_{\nu+1}$, so $T_{\nu+1}$ is normal in T_ν and $T_\nu/T_{\nu+1} = (G_a)^{n-\nu}$. Thus the full triangular group T is solvable.

Now let $S = \{s\}$ be any set of $n \times n$ matrices which commute with each other. Considering S as a set of linear transformations on an n-dimensional vector space, for any $s \in S$ and any quantity α, the null space of $(s - \alpha e)$ is S-invariant. Hence the set of matrices S is reducible (in the sense of representation theory), and by repeated application of this we can reduce all the matrices of S simultaneously to triangular form. Hence every commutative algebraic group of matrices is solvable.

Let the algebraic group G operate on the variety V and let W be the variety of G-orbits on V and τ the natural rational map from V to W. A rational map σ from W into V is called a *cross section* if $\tau\sigma = 1$; this of course implies that if k is a field of definition for G, V, the operation of G on V, W, τ, and σ, and if p is a generic point of W over k, then σp is a point of V at which τ is defined. (Note however that σ need not be defined at all points of W, so σ is a cross section in the topological sense only on an open subset of W.) The most important case is that in which G is connected and V is a principal space with respect to G, i.e., k being as above and $g \times v$ being generic for $G \times V$ over k, then G operates regularly on V and there exists an everywhere defined rational map (defined over k) from the locus over k of $v \times gv$ into G such that $v \times gv \rightarrow g$. (For example, V may be a connected algebraic group and G a connected algebraic subgroup of V operating on V by either of the laws $g(v) = gv$ or $g(v) = vg^{-1}$; W may be taken to be the corresponding homogeneous space.) In this case $p = \tau v$ is

generic for W over k, so σp is defined and hence $k(g) \subset k(\sigma p, g(\sigma p))$. Since $\tau(g(\sigma p)) = p$, we get $k(g(\sigma p)) = k(g, p)$. Clearly the transcendence degree of $k(g, p)$ over k is $\dim G + \dim W = \dim V$, so $g(\sigma p)$ is generic for V over k. Thus V is birationally equivalent over k to $G \times W$, and G operates only on the first factor, there by multiplication on the left.

In the proof of the following lemma we use a number of results in the theory of algebraic curves which may be found in Chevalley's *Introduction to the Theory of Algebraic Functions of One Variable.*

LEMMA. *Let V be a homogeneous space with respect to the algebraic group G $(= G_a$ or $G_m)$ and let k be a field of definition for V and the operation of G on V. Then V has a point that is rational over k.*

If g, p are independent generic points over k of G, V respectively, then gp is generic for V over $k(p)$. In particular, $\dim V \leqq 1$. The case $\dim V = 0$ is trivial, so suppose V is a curve. Since $k(p, gp)$ is a subfield of the purely transcendental extension $k(p, g)$ of $k(p)$, it is an algebraic function field of one variable over the constant field $k(p)$ that has genus zero. But $k(p, gp)$ may be got by taking the function field $k(gp)/k$ and extending the constant field from k to $k(p)$, which is a regular extension of k. Hence $k(gp)/k$ has genus zero; i.e. $k(V)/k$ has genus zero. Therefore V is birationally equivalent over k to a conic C lying in the projective plane. C is complete and both C and V are (absolutely) nonsingular, so the birational correspondence between V and C is biregular between V and $C - S$, where S is a finite subset of C. Thus we may suppose that $V = C - S$. Now any birational transformation on C, in particular one induced by an element of G, has at least one fixed point. Since any nontrivial translation of a homogeneous space with respect to a commutative group is free of fixed points, this implies that S is nonempty. Let D be the projective line containing G $(D = G_a \cup (\infty)$ or $G_m \cup (0) \cup (\infty))$. Then there is an everywhere defined surjective rational map ϕ_p from D to C such that ϕ_p is defined over $k(p)$ and $\phi_p g = gp$. Clearly $S \subset \phi_p(D - G)$. Since the points of S are rational over $k(p)$ and since p could have been taken to be any generic point of V over k, the points of S are rational over k. Since C has genus zero and at least one rational point, it is birationally equivalent over k to a projective line. C thus contains at least three rational points, and since S has at most two, V has a rational point.

Note that the lemma is false if we merely assume that G, an algebraic group defined over k, is biregularly isomorphic (over some extension field of k) to G_a or G_m. For example, let k be the real numbers and let G be the

rotation group in two dimensions operating in the obvious manner on the finite part of the curve $x^2 + y^2 = -1$.

If G is a connected algebraic group possessing a normal chain of connected algebraic subgroups $G = G_0 \supset G_1 \supset \cdots \supset G_r = \{e\}$ such that each G_i $(i = 0, \cdots, r-1)$ is defined over k and possesses a surjective separable rational homomorphism with kernel G_{i+1} to G_a or G_m, the homomorphism also being defined over k, we say that k is a *field of definition for the solvability of* G.

THEOREM 10. *If the connected solvable algebraic group G operates regularly on the variety V and if $\tau\colon V \to W$ is the natural rational map from V to the variety of G-orbits on V, then there exists a cross section $\sigma\colon W \to V$. Furthermore, if k is a field of definition for the solvability of G, for V, the operation of G on V, W, and τ, then σ may be taken to be defined over k.*

In the special case where V is a homogeneous space with respect to G, the theorem merely says that V has a point that is rational over k. We first assume this special case and show that the theorem holds generally. So let G, V, W, τ, k be as above and let v be a generic point of V over k. Then $p = \tau v$ is generic for W over k. By Theorem 2, $k(v)$ is separably generated over $k(p)$ and, since G is connected, $k(p)$ is algebraically closed in $k(v)$. Thus $k(v)$ is a regular extension of $k(p)$, and the locus V_1 of v over $k(p)$ is a variety defined over $k(p)$ that is a closed subset of V. If g is generic for G over $k(v)$, then gv is generic for V over k and $\tau(gv) = p$. Thus the natural k-isomorphism between $k(v)$ and $k(gv)$ is actually a $k(p)$-isomorphism, and so $gv \in V_1$. It follows that G operates regularly on V_1. If v' is any generic point of V_1 over $k(p)$ we have $k(p)(v')$ naturally $k(p)$-isomorphic to $k(p)(v)$, so $k(v')$ is naturally k-isomorphic to $k(v)$; thus v' is a generic point of V over k. Since $\tau v' = \tau v$, Theorem 2 shows that there exists $g' \in G$ such that $g'v = v'$. Since v' could have been taken to be generic for V_1 over $k(v)$, this implies that gv is generic for V_1 over $k(v)$, i.e. V_1 is a prehomogeneous space with respect to G. Thus there exists a homogeneous space birationally equivalent to V_1 over $k(p)$. Since G operates regularly on V_1, the corollary to Theorem 1 shows that this homogeneous space may be taken to be a $k(p)$-open subset V_1' of V_1. Thus there exists a point $p_1 \in V_1'$ that is rational over $k(p)$. Choose $\gamma \in G$ such that $p_1 = \gamma v$. Since $\tau = \tau\gamma^{-1}$, τ is defined at p_1 and $\tau p_1 = \tau v = p$. Thus $\sigma\colon p \to p_1$ gives a cross section that is defined over k. We must now prove our original assumption that if V is homogeneous, then V has a point that is rational over k. Since this is trivial

if $\dim G = 0$ and the lemma if $\dim G = 1$, we assume that $\dim G > 1$ and use induction on $\dim G$. Let G_1 be a connected algebraic subgroup of G having k as a field of definition for its solvability and such that G admits a surjective separable rational homomorphism defined over k and with kernel G_1 to G_a or G_m. Let V' be the variety of G_1-orbits on V and τ' the natural rational map from V to V', both V' and τ' being taken to be defined over k. By Theorem 5 and Corollary, G/G_1 operates on V', this operation being defined over k. If g, v are independent generic points of G, V respectively over k and if ϕ is the natural homomorphism from G to G/G_1, we have $\tau'(gv) = \phi g(\tau' v)$. Since gv is generic for V over $k(v)$, $\phi g(\tau' v)$ is generic for V' over $k(\tau' v)$; i.e. V' is a prehomogeneous space with respect to G/G_1. Thus we may suppose that V' is homogeneous with respect to G/G_1. Then there exists a point $p' \varepsilon V'$ that is rational over k. But V' may also be considered as a homogeneous space with respect to G, so Proposition 1 is applicable to the case τ': $V \to V'$. Let S be the set of all points of V whose τ'-image is p'. Then S is a closed subset of V which, considered as a cycle all of whose coefficients are one, is rational over k. But if $v_1, v_2 \varepsilon V$, then $\tau' v_1 = \tau' v_2$ if and only if $v_2 \varepsilon G_1 v_1$. (This last relation holds when v_1, v_2 are generic for V over k, hence, by transitivity, for all v_1, v_2.) Thus S has only one component, and hence is a variety defined over k. Clearly S is a homogeneous space with respect to G_1. By the induction hypothesis, S, and hence V, has a point that is rational over k.

COROLLARY 1. *If H is a connected solvable algebraic subgroup of the connected algebraic group G, then G is birationally equivalent to $H \times (G/H)$. Furthermore, this birational equivalence may be defined over any field of definition for G, the solvability of H, and the map $G \to G/H$.*

This follows from the discussion immediately preceding the lemma. Note that if σ: $G/H \to G$ is a cross section for the action of H on G, then the birational equivalence is given by $h \times gH \leftrightarrow \sigma(gH)h^{-1}$ or $h \times Hg \leftrightarrow h\sigma(Hg)$ according as G/H is the homogeneous space of left or right cosets of G modulo H.

COROLLARY 2. *If G is a connected solvable algebraic group, then G is rational. More precisely, if k is a field of definition for the solvability of G, then $k(G)$ is a purely transcendental extension of k.*

5. The main structure theorems. An *abelian variety* is a complete connected algebraic group. The basic facts about abelian varieties can be

found in [6]. Since abelian varieties are commutative groups, when there is no danger of confusion we shall denote their group operations additively. The word "abelian" will always refer to abelian varieties, rather than to commutativity.

By a *linear group* we mean an algebraic group which is biregularly isomorphic to an algebraic group of matrices. Note that linear groups, being embeddible in affine spaces, are noncomplete if they have dimension > 0; thus $\{e\}$ is the only algebraic group which is both linear and abelian. Algebraic subgroups and direct products of linear groups (or abelian varieties) are again linear groups (or complete algebraic groups). The groups G_a and G_m are both linear, the former possessing the matrix representation

$$x \to \begin{pmatrix} 1 & x \\ 0 & 1 \end{pmatrix}.$$

If G is an algebraic group and $g \varepsilon G$, we denote the normalizer of g (i.e. the set of all elements of G that commute with g) by N_g. This is clearly an algebraic subgroup of G. The center C of G is the intersection of all N_g's so C is also an algebraic subgroup of G. The component of the identity of an algebraic group G will always be denoted G_0.

THEOREM 11. *Any rational homomorphism of an abelian variety into a linear group, or of a connected linear group into an abelian variety, is trivial.*

Since a rational homomorphic image of an abelian variety is abelian, the first part is proved. So let τ be a rational homomorphism of the connected linear group L into the abelian variety A. τ is trivial if $\dim L = 0$, so suppose $\dim L > 0$ and that our proposition is true for all linear groups of smaller dimension. If L is commutative then according to Section 4 it is solvable, hence rational, hence $\tau L = 0$. If L is not commutative, let k be a field of definition for L, A, and τ and let g be generic for L over k. Then $\dim N_g < \dim L$, so $\tau(N_g)_0 = 0$. Since $g \varepsilon N_g$, τg has some finite order ν. But τg is generic for τL over k, so for each point $p \varepsilon \tau L$ we have $\nu \cdot p = 0$. Since τL is abelian, $\tau L = 0$.

If G is an algebraic group that operates on the variety V, f a rational function on V, and $g \varepsilon G$, we define (as before) the function $\lambda_g f$ by $\lambda_g f(p) = f(g^{-1}p)$. In the following, we consider vector spaces of functions on V over constants, i.e. over the universal domain. In particular, we consider finite dimensional vector spaces S of functions on V with the property that $\lambda_g S = S$ for all $g \varepsilon G$. If V is a homogeneous space with respect to G

then any such space S consists of functions that are finite everywhere on V, for if W is a proper closed subset of V such that each function of S is finite on $V-W$, then each function in $S=\lambda_g S$ is finite on $V-gW$ for all $g\,\varepsilon\, G$. We need another definition: If the algebraic group G operates on V, there may exist a surjective rational homomorphism τ from G to an algebraic group G' that also operates on V, and in such a way that $gp=(\tau g)p$; if any such τ is necessarily a biregular isomorphism, we say that G operates *faithfully* on V. By [4, Proposition 2], this is equivalent to the definition given by Weil when G is connected. Clearly the operation of any connected algebraic group on itself by left translation is faithful.

THEOREM 12. *Let the algebraic group G operate regularly on the non-singular variety V, all defined over the field k. Then any everywhere finite rational function on V is contained in a finite dimensional vector space S of such functions such that $\lambda_g S=S$ for all $g\,\varepsilon\, G$, and S may be taken so as to have a basis consisting of functions that are defined over k. If λ_g induces the linear transformation Λ_g on such a space S, then (choosing such a basis for S) the map $g\rightarrow\Lambda_g$ is a surjective rational homomorphism defined over k from G to an algebraic group of matrices. If G operates faithfully on V and the functions of S generate the entire function field of V, then this homomorphism is a biregular isomorphism.*

By [3, Theorem 10, p. 239], any everywhere finite rational function f on V can be written $f=\sum_\nu \alpha_\nu\phi_\nu$, where the α_ν's are constants and each ϕ_ν is an everywhere finite function in $k(V)$, so to prove that f is contained in some space S such as above it suffices to assume that $f\,\varepsilon\, k(V)$. This being so, let g be a generic point over k for some component G_α of G. Then $\lambda_g f\,\varepsilon\, k(g)(V)$ is everywhere finite on V, so (again by [3, Theorem 10, p. 239]) we can write $\lambda_g f=\sum_{i=1}^N c_i(g)\phi_i$, where each $c_i(g)\,\varepsilon\, k(g)$, each $\phi_i\,\varepsilon\, k(V)$ is everywhere finite, and $c_1(g),\cdots,c_N(g)$ are linearly independent over k. Let $g_1,\cdots,g_N$ be independent generic points of G_α over k and consider the $N\times N$ matrix $(c_i(g_j))$. If we had $|c_i(g_j)|=0$, then $c_1(g_1),\cdots,c_N(g_1)$ would be linearly dependent over $k(g_2,\cdots,g_N)$, contradicting the linear disjointness of $k(g_1)$ and $k(g_2,\cdots,g_N)$ over k and the linear independence over k of $c_1(g_1),\cdots,c_N(g_1)$; thus $|c_i(g_j)|\neq 0$. Since $\lambda_{g_j}f=\sum_{i=1}^N c_i(g_j)\phi_i$, for $j=1,\cdots,N$, we can write $\phi_i=\sum_{j=1}^N d_j\lambda_{g_j}f$, where $d_1,\cdots,d_N\,\varepsilon\, k(g_1,\cdots,g_N)$. Let S be the finite dimensional vector space generated by all the ϕ_i's we obtain by letting g range over a set of generic points over k of the various components of G. If g' is generic over k for any component G_β of G, then the $g_1,\cdots,g_N$ used above could have been taken to be independent generic points over $k(g')$

of G_a, whence $\lambda_{g'}\phi_i = \sum_j d_j \lambda_{g'g_j} f \varepsilon S$, since each $g'g_j$ is generic for a component of G over k. Thus $\lambda_{g'}S \subset S$, whence $\lambda_{g'}S = S$. If γ is any point of G, we can write $\gamma = g'g''$, where g', g'' are generic points over k of components of G, so $\lambda_\gamma S = \lambda_{g'g''}S = \lambda_{g'}\lambda_{g''}S = \lambda_{g'}S = S$. Clearly $f \varepsilon S$, so S is the space we were looking for. Now let $f_1, \cdots, f_n$ be everywhere finite functions on V that are a basis for a finite dimensional invariant space S and suppose that each $f_i \varepsilon k(V)$. For any $g \varepsilon G$ we can write $\lambda_g f_i = \sum_j c_j{}^i(g) f_j$, where the $c_j{}^i(g)$'s are well-determined constants depending on g. We now show that each $c_j{}^i(g) \varepsilon k(g)$. Since $\lambda_g f_i$ is everywhere finite and defined over $k(g)$, we can write $\lambda_g f_i = \sum_{\nu=1}{}^N a_\nu{}^i(g)\psi_\nu$, where each $a_\nu{}^i(g) \varepsilon k(g)$, each $\psi_\nu \varepsilon k(V)$, and where $\psi_1, \cdots, \psi_N$ are linearly independent over k. Consideration of the vector space over k generated by $f_1, \cdots, f_n, \psi_1, \cdots, \psi_N$ shows that if we alter the f_i's and the ψ_ν's by suitable linear transformations with coefficients in k we can obtain $f_i = \psi_i$ for $i < s (1 \leqq s \leqq n+1)$ while $f_1, \cdots, f_n, \psi_s, \cdots, \psi_N$ are linearly independent over k. Then the equation $\sum_j c_j{}^i(g) f_j = \sum_\nu a_\nu{}^i(g)\psi_\nu$ gives

$$\sum_{j<s} (c_j{}^i(g) - a_j{}^i(g)) f_j + \sum_{j \geqq s} c_j{}^i(g) f_j = \sum_{\nu \geqq s} a_\nu{}^i(g)\psi_\nu,$$

and therefore $c_j{}^i(g) = a_j{}^i(g)$ for $j < s$, $c_j{}^i(g) = 0$ for $j \geqq s$. Thus each $c_j{}^i(g) \varepsilon k(g)$. The matrix $\Lambda_g = (c_j{}^i(g))$ is invertible (since it has the inverse $\Lambda_{g^{-1}}$) so $g \to \Lambda_g$ is a homomorphism of G into the group of $n \times n$ invertible matrices. By Theorem 3 (first applied to G_0), this is a surjective rational homomorphism from G to an algebraic group of matrices, and the homomorphism and matrix group are defined over k. In particular, the $c_j{}^i$'s are everywhere finite rational functions on the various components of G. In the case where $k(f_1, \cdots, f_n) = k(V)$, clearly the matrix group itself operates on V, so if G operates faithfully on V, then the homomorphism from G to the matrix group is a biregular isomorphism.

Corollary 1. *If the algebraic group G is such that G_0 is linear, then G is linear. If G is linear and defined over k, then G admits a biregular isomorphism defined over k to an algebraic group of matrices (which is therefore also defined over k).*

Let G be defined over k and let G_0 be linear. If G is finite, we may use the regular representation of G, so assume $\dim G > 0$. Let V be the direct product of the various components of G, each taken once. V is non-singular and defined over k. Left translation by elements of G defines an operation of G on V (in which any $g \varepsilon G_0$ operates on each of the direct factors of V, while if $g \varepsilon G$, $g \not\varepsilon G_0$, then g permutes the various direct factors),

and this operation is clearly defined over k, regular, and faithful. Since G_0 is linear, the coordinate functions in a matrix representation of G_0 give a set of everywhere finite rational functions on G_0 which generate the entire function field of G_0. Hence for each component G_α of G we can find a finite set of everywhere finite rational functions on G_α which generate the entire function field of G_α. Thus there exists a finite set of everywhere finite rational functions on V which generate the entire function field of V. Now apply the theorem.

COROLLARY 2. *If there exists a surjective rational homomorphism from the algebraic group G to a linear group of the same dimension, then G is linear.*

Let $\tau\colon G \rightarrow H$ be the rational homomorphism in question. We may suppose that G, and therefore also H, is connected. Letting $f_1, \cdots, f_n$ be a set of everywhere finite rational functions on H that generate the entire function field of H and letting k be a field of definition for G, H, τ, and each f_i, we get $k(H) = k(f_1, \cdots, f_n)$. Each f_i induces under τ^{-1} an everywhere finite function in $k(G)$, which we also denote f_i. Since $k(G)$ is a finite algebraic extension of $k(H)$, we can find elements $F_1, \cdots, F_s \varepsilon k(G)$ that are integrally dependent on $k[f_1, \cdots, f_n]$ such that $k(G) = k(f_1, \cdots, f_n, F_1, \cdots, F_s)$. Since each F_j is everywhere finite on G, we can apply the theorem to G operating on itself by left translation.

COROLLARY 3. *If G is an algebraic group that is defined over k, then there exists a connected normal algebraic subgroup D of G, also defined over k, such that G/D is linear and such that the kernel of any rational homomorphism from G to a linear group contains D.*

If D_1 is the kernel of a rational homomorphism from G to a linear group, Corollary 2 shows that G/D_1 is linear. If D_2 is another normal algebraic subgroup of G such that G/D_2 is linear, then the kernel of the obvious rational homomorphism from G into the linear group $(G/D_1)\times(G/D_2)$ is $D_1 \cap D_2$, so $G/(D_1 \cap D_2)$ is linear. Hence there exists a smallest normal algebraic subgroup D of G such that G/D is linear. By Corollary 2, G/D_0 is linear, so D is connected. It remains to show that D is defined over k, and for this we may suppose G connected. Let $\tau\colon G \rightarrow G/D$ be the natural homomorphism, let G and τG $(= G/D)$ operate on themselves by left translation, and let S be a finite dimensional invariant vector space of everywhere finite rational functions on τG that generates the entire function field of τG. If $g \varepsilon G$, then λ_g, $\lambda_{\tau g}$ are automorphisms of the function fields of G, τG respectively, and for any $f \varepsilon S$ we have $f\tau$ everywhere finite on G and $\lambda_g(f\tau) = (\lambda_{\tau g}f)\tau$.

Hence S_τ is a finite dimensional vector space of everywhere finite rational functions on G that is invariant under each λ_g. If Λ_g, $\Lambda_{\tau g}$ are the linear transformations induced by λ_g, $\lambda_{\tau g}$ respectively on S_τ, S, and if we choose corresponding bases for S_τ and S, then we have a matrix equality $\Lambda_g = \Lambda_{\tau g}$. If $S' \supset S_\tau$ is a finite dimensional invariant vector space of everywhere finite rational functions on G and Λ'_g is the linear transformation on S' induced by λ_g, then the map $g \to \Lambda'_g$ is a surjective rational homomorphism from G to a linear group G'. But $\tau g \to \Lambda_{\tau g}$ is a biregular isomorphism and Λ_g is merely the restriction of Λ'_g to S_τ, so we have a sequence of rational homomorphisms $G \to G' \to \tau G$. Since G' is linear, the kernel of $G \to G'$ must be D, so $G' \to \tau G$ is an isomorphism. But τ is separable, so G' is biregularly isomorphic to τG. Now note that S' could have been taken to have a basis consisting of functions in $k(G)$, in which case the map $G \to G'$ can be taken to be defined over k. This map being separable, Proposition 1, Corollary shows that its kernel D is a rational cycle over k. Hence D is defined over k.

The above proof also shows that if G is connected and $\tau\colon G \to G/D$ is taken to be defined over k then, under the obvious identification, the set of everywhere finite functions is $k(G/D)$ is precisely the set of everywhere finite functions in $k(G)$. For it is clear that any such function on G/D is one on G, so let $f \in k(G)$ be everywhere finite on G. The space S' used in the proof could have been taken so large as to include f. Letting $f_1, \cdots, f_n \in k(G)$ be a basis for S' and writing $\lambda_g f_i = \sum_j c_j{}^i(g) f_j$, the proof shows that the $c_j{}^i$'s are everywhere finite functions on G/D, and one merely has to note that $f_i(g) = (\lambda_{g^{-1}} f)(e) = \sum_j c_j{}^i(g^{-1}) f_j(e) \in k(\tau g)$.

THEOREM 13. *Let G be a connected algebraic group defined over k and let C be its center. Then C is a k-closed normal algebraic subgroup of G and G/C is linear.*

C is known to be an algebraic subgroup of G; since it is invariant with respect to all k-automorphisms of the universal domain, C is k-closed. Let $\mathfrak{o}$, a subring of the field of all rational functions on G, be the local ring of e, and let $\mathfrak{m}$ be its maximal ideal. Let $f_1, \cdots, f_n \in \mathfrak{m} \cap k(G)$ be a set of uniformizing parameters at e. For any rational function f on G and any $g \in G$ define the function $\omega_g f$ by $\omega_g f(p) = f(g^{-1}pg)$. If $g_1, g_2 \in G$, then $\omega_{g_1}\omega_{g_2} = \omega_{g_1 g_2}$, so the map $g \to \omega_g$ is a homomorphism from G to a group of automorphisms of the function field of G leaving constants fixed. For any $g \in G$, $\omega_g \mathfrak{o} = \mathfrak{o}$, so $\omega_g \mathfrak{m}^\nu = \mathfrak{m}^\nu$ for any integer $\nu > 0$. $\mathfrak{o}/\mathfrak{m}$ is the field of constants, so $\mathfrak{m}/\mathfrak{m}^\nu$ is a vector space over the constants having as a basis the various elements $\bar{f}_1{}^{i_1} \cdots \bar{f}_n{}^{i_n}$, where $i_1, \cdots, i_n$ are integers $\geqq 0$ of positive sum $< \nu$ and where

$\bar{f}$ denotes the residue class of a function $f \varepsilon \mathfrak{m}$ in the natural map $\mathfrak{m} \to \mathfrak{m}/\mathfrak{m}^\nu$. For any $g \varepsilon G$, ω_g induces an invertible linear transformation $\bar{\omega}_g$ on $\mathfrak{m}/\mathfrak{m}^\nu$. Considering the action of ω_g on $\mathfrak{o} \cap k(g)(G)$ shows that $\bar{\omega}_g \bar{f}_i$ is of the form

$$\bar{\omega}_g \bar{f}_i = \sum_{i_1 + \cdots + i_n < \nu} c^{(i)}{}_{i_1 \cdots i_n}(g) \bar{f}_1{}^{i_1} \cdots \bar{f}_n{}^{i_n},$$

where each $c^{(i)}{}_{i_1 \cdots i_n}(g) \varepsilon k(g)$. By Theorem 3, the map τ: $g \to \bar{\omega}_g$ is a rational homomorphism defined over k of G into a group of invertible square matrices, and hence τG is an algebraic group of matrices. In particular, each $c^{(i)}{}_{i_1 \cdots i_n}$ is an everywhere defined rational function on G. The kernel of τ clearly contains C. If $\gamma \varepsilon G - C$, then there exists a function $f \varepsilon \mathfrak{m}$ such that $\omega_\gamma f \neq f$, so for ν sufficiently large we have $\bar{\omega}_\gamma \bar{f} \neq \bar{f}$. Thus, for ν sufficiently large, $\bar{\omega}_g = 1$ only if g is on a closed subset of G not passing through γ. This being true for each $\gamma \varepsilon G - C$, for sufficiently large ν the kernel of τ is precisely C. In this case we have a rational isomorphism from G/C to the linear group τG. By the previous Corollary 2, G/C is linear.

COROLLARY 1. *If G is connected and D is as in the previous Corollary* 3, *then $D \subset C$.*

COROLLARY 2. *If the connected algebraic group G has a commutative normal algebraic subgroup H such that G/H is an abelian variety, then G is commutative.*

We have a natural surjective rational homomorphism from G/H to G/HC, so G/HC is an abelian variety. Since we also have a surjective rational homomorphism from the linear group G/C to G/HC, the latter consists of only one element. Thus $G = HC$. That is, $G = C$.

PROPOSITION 3. *Let τ be a rational map of the connected algebraic group G into the abelian variety A such that $\tau e = 0$. Then τ is a homomorphism.*

Consider the rational map of $G \times G$ into A defined by $g_1 \times g_2 \to \tau(g_1 g_2)$. By [6, Cor., p. 32] we can write $\tau(g_1 g_2) = \tau_1 g_1 + \tau_2 g_2$, where τ_1, τ_2 are rational maps of G into A. By [6, Theorem 6, p. 27], τ, τ_1, τ_2 are everywhere defined. Assuming, as we may, that $\tau_1 e = 0$, we get $\tau_2 e = \tau e = 0$. Thus $\tau_1 g_1 = \tau(g_1 e) = \tau g_1$, and similarly $\tau_2 g_2 = \tau g_2$. That is, $\tau(g_1 g_2) = \tau g_1 + \tau g_2$.

The following is a slight generalization of Weil's abstract analogue of the Poincaré theorem on complete reducibility [6, Theorem 26, p. 94], itself contained implicitly in [5, § 5].

THEOREM 14. *Let G be a connected commutative algebraic group and V a principal space with respect to G, all defined over k. Then there exists a rational map $\phi: V \to G$ such that ϕ is defined over k and if g, v are independent generic points of G, V respectively over k, then $\phi(gv) = \phi(v) + n \cdot g$, where n is an integer $\neq 0$.*

Here G operates regularly on V, and if k, g, v are as above and T is the locus on $V \times V$ of $v \times gv$ over k, then there is an everywhere defined rational map $\theta: T \to G$ such that θ is defined over k and $\theta(v \times gv) = g$. If $p \varepsilon V$ and $p' \varepsilon V$ is in the closure of the orbit Gp, then $p \times p' \varepsilon T$, so $p' = (\theta(p \times p'))p$; hence all orbits on V are closed. Thus the orbit Gv is precisely the locus on V of gv over $k(v)$. Let τ be the natural rational map from V to its variety of G-orbits, τ being taken to be defined over k. Then $k(v)$ is a regular extension of $k(\tau v)$, so let V' be the locus on V of v over $k(\tau v)$. Since $\tau v = \tau(gv)$, the natural k-isomorphism of $k(v)$ and $k(gv)$ is also a $k(\tau v)$-isomorphism, so V' is also the locus of gv over $k(\tau v)$. Hence $V' \supset Gv$. On the other hand, if v' is any generic point of V' over $k(\tau v)$, then $v' \varepsilon Gv$ (for, by a dimension argument, v' is generic for V over k and $\tau v' = \tau v$). Thus $V' = Gv$; in particular, Gv is defined over $k(\tau v)$. If $\sum_{i=1}^{n} x_i$ is a positive zero-cycle on Gv that is rational over $k(\tau v)$, then for each i we have $v \times x_i \varepsilon T$, so we can form $\sum_i \theta(v \times x_i)$, the sum being taken in G. Using the commutativity of G, the main theorem on symmetric functions [6, Theorem 1, p. 15] shows that $\phi(v) = \sum_i \theta(v \times x_i)$ is rational over $k(v)$. Since $\tau(gv) = \tau v$, we get

$$\phi(gv) = \sum_i \theta(gv \times x_i) = \sum_i \theta(v \times x_i) - n \cdot g = \phi(v) - n \cdot g.$$

COROLLARY. *Let the abelian variety A be an algebraic subgroup of the connected algebraic group G, both defined over k. Then A is contained in the central subgroup D of G and there exists a connected k-closed algebraic subgroup G_1 of G such that $G = G_1 A$ and $G_1 \cap A$ is finite. If G is non-complete, so is G_1.*

In the natural homomorphism from G to the linear group G/D, A must go into the identity, so $A \subset D$. Applying the theorem to the case where G, V are replaced by A, G (A operating on G by left translation) we get a rational map $\phi: G \to A$ such that ϕ is defined over k and such that for $g \varepsilon G$, $x \varepsilon A$ we have $\phi(xg) = \phi(g)x^n$, $n \neq 0$. ϕ is everywhere defined, so we may alter it by a translation on A to get $\phi(e) = e$, in which case ϕ is a homomorphism. Since $\phi(x) = x^n$, ϕ is surjective and its kernel meets A in only a finite number of points. Letting G_1 be the component of the identity of the

kernel of ϕ, we get G_1 k-closed, $G_1 \cap A$ finite, and $\dim G = \dim G_1 + \dim A$. Since the kernel of the natural homomorphism from $G_1 \times A$ to G is finite, we get $G = G_1 A$. Finally, if G_1 is complete, so is $G_1 \times A$, hence also G.

The given form of the following lemma is due to Chow.

LEMMA 1. *Let V be an abstract variety defined over k. Then there exists a projective variety V', a birational map τ from V' to V, both V' and τ being defined over k, and a k-closed subset F of V' such that τ is everywhere defined on $V' - F$ and no point of F corresponds under τ to a point of V. Considering τ as a set-theoretic map, the image of $V' - F$ is V and the inverse image of any point of V is a closed subset of V' that is disjoint from F.*

Let V be given by a coherent set of birational correspondences among $V_1 - F_1, \cdots, V_n - F_n$, where each V_i is a projective variety and F_i is a frontier on V_i, everything being defined over k. Let $P_1, \cdots, P_n$ be corresponding generic points of $V_1, \cdots, V_n$ respectively over k, let V' be the projective variety which is the locus of $P_1 \times \cdots \times P_n$ over k, and let τ be the natural birational map from V' to V. If τ_i is the birational map from V' to V_i defined by $\tau_i(P_1 \times \cdots \times P_n) = P_i$, then τ_i is everywhere defined. Let $F_i' = \tau_i^{-1}\{F_i\}$ and set $F = F_1' \cap \cdots \cap F_n'$. If $p \varepsilon F$, then $\tau_i p \varepsilon F_i$, so p corresponds to no point of V. On the other hand, if $p \varepsilon V' - F$ then for some i we have $p \varepsilon V' - F_i'$, so τ_i is defined at p and $\tau_i p \varepsilon V_i - F_i$; thus τ is defined at p. The last statement follows from the fact that each point of V corresponds to a nonempty closed subset of V'.

LEMMA 2. *Let V^n be a projective variety defined over k and let F be a nonempty k-closed proper subset of V. Then there exists a projective variety V' and a birational map $\tau\colon V' \to V$, all defined over k, such that τ is everywhere defined on V' and* $\dim \tau^{-1}\{F\} = n - 1$.

This is contained in [7]: Since it may be replaced by a derived normal model, V may be supposed relatively normal with reference to k. If $\dim F < n - 1$, let V_1 be the monoidal transform of V with respect to F and τ_1 the birational map from V_1 to V. Then τ_1 is everywhere defined and $\dim \tau_1^{-1}\{F\} > \dim F$ ([7, pp. 533, 520]). If necessary repeat this process.

LEMMA 3. *Let U^m be a variety, V^n a projective variety, $\phi\colon U \to V$ a generically surjective rational map, and X^{m-1} a simple subvariety of U, all defined over k. Then there exists a projective variety V', birationally equivalent over k to V, such that the map from V' to V is everywhere defined and*

such that if $\phi'\colon U \to V'$ *is the map corresponding to* ϕ *and* x *is generic for* X *over* k, *then the locus over* k *of* $\phi' x$ *has dimension* $\geqq n-1$.

The map ϕ enables us to identify $k(V)$ with a subfield of $k(U)$. Let θ be the valuation of $k(U)$ (with value group the ordinary integers) given by the order of vanishing of functions on U along X. Then the residue field with respect to θ of $k(U)$ is naturally k-isomorphic to $k(X)$, hence has transcendence degree over k equal to $m-1$. Now the transcendence degree of $k(U)$ over $k(V)$ is $m.-n$, so the residue field of $k(U)$ with respect to θ has a transcendence degree over the residue field of $k(V)$ with respect to the valuation induced on it by θ which is $\leqq m-n$. Thus the residue field of $k(V)$ with respect to θ has transcendence degree $\geqq (m-1)-(m-n) = n-1$ over k. Hence there exist $f_1, \cdots, f_{n-1} \varepsilon k(V)$ which, considered as functions on U, are finite along X and induce functions on X which are algebraically independent over k; in particular, $f_1, \cdots, f_{n-1}$ are defined at x and assume at this point values $\xi_1, \cdots, \xi_{n-1}$ which are algebraically independent over k. Let V' be the graph of the rational map from V to the projective space of dimension $n-1$ determined by $(1, f_1, \cdots, f_{n-1})$. The birational map from V' to V is defined over k and everywhere defined on V'. Since V' is a subvariety of the direct product of two projective spaces, it is itself a projective variety. Since X^{m-1} is simple on U and V' is complete, $\phi' x$ is defined. The lemma now follows from the fact that $\xi_1, \cdots, \xi_{n-1} \varepsilon k(\phi' x)$.

THEOREM 15. *Let the connected algebraic group* G *operate regularly on the noncomplete variety* V^n, *all defined over the algebraically closed field* k. *Then there exists a normal projective variety* V', *defined and birationally equivalent to* V *over* k, *and a subvariety* W^{n-1} *of* V', *also defined over* k, *such that if we consider the operation of* G *on* V' *that is induced by its operation on* V *and let* g, x *be independent generic points of* G *and* W *respectively over* k, *then* gx *is defined and the rational map defined over* k *by* $g \times x \to gx$ *defines an operation of* G *on* W.

Let V', F, τ be as in Lemma 1; here F is nonempty. Note that if V'' is a projective variety and $\psi\colon V'' \to V'$ a birational map defined over k that is everywhere defined on V'', then V'', $F' = \psi^{-1}\{F\}$, and $\tau' = \tau\psi$ also satisfy the conclusions of Lemma 1. Hence, by Lemma 2, we may suppose that $\dim F = n-1$. Replacing V' by a derived normal model if necessary, we may suppose V' normal with reference to k, hence (absolutely) normal. In what follows, we use repeatedly the facts that a normal variety has no singular subvarieties of codimension one and that a rational map from a variety to a complete variety is defined along any simple subvariety of codimension one.

Let W_0 be a component of F of dimension $n-1$. The operation of G on V induces an operation of G on V', so apply Lemma 3 to the rational map $\phi\colon G\times V'\to V'$ defined by $\phi(g\times v)=gv$ and the subvariety $G\times W_0$ of $G\times V'$. We get a variety V'' and an everywhere defined rational map $\rho\colon V''\to V'$ (V'' and ρ being defined over k) such that if g, p are independent generic points of G, W_0 respectively over k, then $\rho^{-1}\phi$ is defined at $g\times p$ and $\dim_k \rho^{-1}\phi(g\times p)\geqq n-1$. Replacing V'' by a derived normal model if necessary, we may suppose V'' normal. Now the birational correspondence between V' and V'' is biregular between p and $\rho^{-1}p$, and that between $G\times V'$ and $G\times V''$ is biregular between $g\times p$ and $g\times\rho^{-1}p$. Replacing V' and W_0 by V'' and the locus over k of $\rho^{-1}p$, we have the following situation: V', F, τ satisfy the conclusions of Lemma 1, V' is normal, and F has a component W_0 of dimension $n-1$ such that if g, p are independent generic points of G, W_0 respectively over k, then gp (which is defined) has a locus W over k which has dimension $\geqq n-1$. If $gp\not\varepsilon F$, then $\tau(gp)\,\varepsilon\, V$, so p corresponds under τ to the point $g^{-1}\tau(gp)\,\varepsilon\, V$, which is false. Hence $gp\,\varepsilon\, F$, implying $\dim W=n-1$. We claim that our present V' and W satisfy the demands of the theorem. For let g_1 be a generic point of G over $k(g,p)$. Then $g_1(gp)$ is defined, hence (by the lemma to Theorem 1) so is $(g_1g)p$ and they are equal. $(g_1g)p$ is generic for W over k; in particular the map $g_1\times gp\to g_1(gp)$ defines a rational map (defined over k) from $G\times W$ to W. Clearly $g_1(gp)$ is rational over $k(g_1,gp)$. But g_1^{-1} is generic for G over $k(g_1g,p)$, so $g_1^{-1}(g_1(gp))$ is defined and equals gp. Thus $k(g_1,g_1(gp)) =k(g_1,gp)$. If now g_2 is generic for G over $k(g_1,g,p)$, then clearly $g_1(g_2(gp))=g_1g_2gp=(g_1g_2)(gp)$, completing the proof. We may add that for any $\gamma\,\varepsilon\, G$, the birational map T_γ on V' given by $T_\gamma v=\gamma v$ may be applied to a generic point of W over $k(\gamma)$ and then induces a birational map on W which is the same as that produced by γ in the operation of G on W; for if g, p are independent generic points of G, W_0 respectively over $k(\gamma)$, then $T_\gamma(gp)$ is defined and (as a specialization of the relation $T_\gamma(gv)=(\gamma g)v$) has the value $(\gamma g)p$ while $\gamma(gp)$ (defined according to the operation of G on W) equals $(\gamma g_1^{-1})(g_1(gp))$ (g_1 being generic for G over $k(\gamma,g,p)$), which equals $(\gamma g_1^{-1})((g_1g)p)=(\gamma g)p$, by Theorem 1, Lemma.

LEMMA 1. *Any noncomplete algebraic group has an algebraic subgroup of dimension >0 which is linear.*

If the noncomplete algebraic group G has dimension one, then G itself is linear, so we use induction on $\dim G$. We may assume G connected. Since G operates regularly on itself by left translation, Theorem 15 gives us

the existence of a normal variety V which is birationally equivalent to G and a subvariety W of V of dimension one less than G such that the operation of G on V induces an operation of G on W. Let W' be a variety birationally equivalent to W such that G operates regularly on W'. Fix a point $P \varepsilon W'$ and let H_P be the algebraic subgroup of G consisting of all g such that $gP = P$. The map $g \rightarrow gP$ is an everywhere defined rational map of G into W', and for any $\gamma \varepsilon G$ the points of G that map into γP are precisely γH_P. Since $\dim GP \leqq \dim W' < \dim G$, we get $\dim H_P > 0$. First assume that $H_P \neq G$. If H_P is noncomplete we can apply our induction assumption to H_P; if H_P is complete, Theorem 14, Corollary shows that G has another proper algebraic subgroup which is noncomplete, and we again use our induction assumption. We are thus reduced to the case where $H_P = G$ for all $P \varepsilon W'$, i. e. where the operation of G on W induced by its operation on V is trivial. Let k be an algebraically closed field of definition for G, V, the operation of G on V, and W. If g, P are independent generic points over k of G, W respectively, then the operation of G on V is defined at the point $g \times P \varepsilon G \times V$ and $gP = P$. Since V is normal, W is a simple subvariety of V. Thus we can find a point $p \varepsilon W$, p rational over k, such that p is simple on V and gp is defined (according to the operation of G on V) and equals p whenever g is generic for G over k. But any point of G can be expressed as the product of two points that are generic for G over k. Thus gp is defined and equals p for *all* $g \varepsilon G$. Let $\mathfrak{o}$ be the local ring of p in the function field of V (isomorphic to that of G), $\mathfrak{m}$ the maximal ideal of $\mathfrak{o}$. For any $g \varepsilon G$ the previously defined operator λ_g satisfies the relation $\lambda_g \mathfrak{o} = \mathfrak{o}$; therefore $\lambda_g \mathfrak{m}^\nu = \mathfrak{m}^\nu$ for each integer $\nu > 0$. The method of proof of Theorem 13 can now be applied to the present case. Since for each $g \varepsilon G$, $g \neq e$, one can find a function $f \varepsilon \mathfrak{m}$ such that $\lambda_g f \neq f$, this proof shows that we have a rational isomorphism from G into an algebraic group of matrices. Thus G itself is linear.

LEMMA 2. *If the connected algebraic group G possesses a normal algebraic subgroup H which is biregularly isomorphic to G_a or G_m and such that G/H is linear, then G is linear.*

Let τ be the natural map from G to G/H. H operates on G by the rule $h(g) = hg$, so by Theorem 10 there exists a rational map σ: $G/H \rightarrow G$ such that $\tau\sigma = 1$. For any $g \varepsilon G$ such that σ is defined at τg we have $\tau((\sigma\tau g)g^{-1}) = e$, so $(\sigma\tau g)g^{-1} \varepsilon H$. Let α be a coordinate function on H, got from the obvious coordinate function on G_a or G_m; then α is everywhere defined and finite on H and $\alpha(h_1) = \alpha(h_2)$ if and only if $h_1 = h_2$. The function $\alpha((\sigma\tau g)g^{-1})$ is a rational function on G that is defined and finite

everywhere outside $\tau^{-1}(W)$, where W is the closed subset of G/H on which σ is not defined. Since G/H is linear, we can find a nonzero everywhere finite rational function w on G/H that vanishes on W. Considering w as a function on G, w vanishes on $\tau^{-1}(W)$. Hence for ν sufficiently large the rational function f on G defined by $f(g) = w^\nu \alpha((\sigma\tau g)g^{-1})$ is everywhere finite. Consider the operation of G on itself by the rule $g(p) = gp$. By Theorem 12, f is contained in a finite dimensional vector space S of rational functions on G such that $\lambda_g S = S$ for each $g \varepsilon G$, and the action of λ_g on S induces a rational homomorphism τ' from G to a linear group G'. If $h \varepsilon H$, $h \neq e$, we have

$$\begin{aligned}\lambda_h f(g) &= f(h^{-1}g)\\ &= w^\nu(h^{-1}g)\alpha((\sigma\tau(h^{-1}g))g^{-1}h) = w^\nu(g)\alpha((\sigma\tau g)g^{-1}h) \neq f(g).\end{aligned}$$

Thus the kernel of τ' meets H only in the point e. Hence we obtain a rational isomorphism from G into the linear group $(G/H)\times G'$. By Theorem 12, Corollary 2, G is linear.

THEOREM 16 (Chevalley). *Let G be a connected algebraic group. Then there exists a linear connected normal algebraic subgroup L of G such that G/L is an abelian variety. L is unique and contains all other connected linear algebraic subgroups of G. If G is defined over k, then L is k-closed.*

First assume the first contention. Then if L' is any connected linear algebraic subgroup of G, it must map into 0 in the natural homomorphism from G to the abelian variety G/L, so $L' \subset L$. This also proves the unicity of L. If G is defined over k, then L is invariant with respect to all k-automorphisms of the universal domain, hence it is k-closed. It remains to prove the main assertion, and here we use induction on $\dim G$. The case $\dim G = 0$ is trivial, so assume $\dim G > 0$ and that our contention is true for all connected algebraic groups of dimension smaller than G. If the center C of G is complete then C_0 is an abelian variety, so by Theorem 14, Corollary there exists a connected algebraic subgroup G_1 of G such that $G = G_1C_0$ and $G_1 \cap C_0$ is finite. G_1 is a normal subgroup of G and its center is finite, so G_1 is linear. We also have a natural surjective rational isomorphism from $C_0/(C_0 \cap G_1)$ (which is an abelian variety) to $G_1C_0/G_1 = G/G_1$, so G/G_1 is an abelian variety. We must now consider the case in which C is noncomplete. Here Lemma 1 gives us the existence of a linear algebraic subgroup H of C such that $\dim H > 0$. Since H is linear and commutative, it is solvable, hence contains an algebraic subgroup which is biregularly isomorphic to G_a or G_m. Hence we may assume to begin with that H is a normal algebraic subgroup of G which is biregularly isomorphic to G_a or G_m. By our induction

assumption G/H has a linear connected normal algebraic subgroup, which we may write L/H (L being a connected normal algebraic subgroup of G that contains H), such that $(G/H)/(L/H)$ is abelian. Since this latter algebraic group is biregularly isomorphic to G/L, G/L is an abelian variety. Since both H and L/H are linear, Lemma 2 implies the linearity of L.

We remark that if the algebraic group G is defined over a field k of characteristic $p \neq 0$ then L need not be defined over k.

COROLLARY 1. *If the algebraic group G possesses a rational homomorphism whose kernel is linear (complete) into a linear (complete) algebraic group, then G is linear (complete).*

Let τ: $G \rightarrow G'$ be the rational homomorphism in question, and let H be the kernel of τ. We may suppose that G is connected. Let L be the maximal connected linear algebraic subgroup of G and first suppose that H and G' are both linear. Since the kernel of τ contains H_0, there is a natural rational homomorphism from G/H_0 into G'; since the kernel of this homomorphism is the finite group H/H_0, G/H_0 must be linear. Since $H_0 \subset L$, there is a natural surjective rational homomorphism from the linear group G/H_0 to the abelian variety G/L. Thus G/L consists of only one point, i.e. $G = L$. Finally suppose that H and G' are complete. Then we must have $\tau L = 0$, so $L \subset H$. Since H is complete, we get $L = \{e\}$, so $G = G/L$ is an abelian variety.

COROLLARY 2. *If there exists a surjective rational homomorphism from the linear (or complete) algebraic group G to the algebraic group H, then H is linear (or complete).*

This is trivial in the complete case, so suppose G linear. Then any nontrivial rational homomorphism from H_0 to an abelian variety would induce one of G_0.

COROLLARY 3. *If G_1 and G_2 are isogenous connected algebraic groups and if either G_1 or G_2 is linear (or abelian) then so is the other.*

COROLLARY 4. *Any solvable algebraic group is linear.*

Note that L is the smallest normal algebraic subgroup of the connected algebraic group G giving rise to an abelian factor group (for if G/L' is abelian, then the natural rational homomorphism from L into G/L' is trivial, so $L \subset L'$). Theorem 12, Corollary 3 gives the analogous result for linear factor groups of G.

COROLLARY 5. ***Let G be a connected algebraic group, let L be its maximal connected linear algebraic subgroup, and let D be the smallest normal algebraic***

subgroup of G giving rise to a linear factor group. Then $G = LD$, and if D' is any algebraic subgroup of G such that $G = LD'$, then $D' \supset D$. Any rational homomorphism from D to a linear group is trivial, and D contains only a finite number of elements of any given finite order.

G/LD is a rational homomorphic image of both the linear group G/D and the abelian variety G/L, so $G/LD = \{e\}$; i.e. $G = LD$. If D' is any algebraic subgroup of G such that $G = LD'$, to prove that $D' \supset D$ it suffices to take D' connected. Then $D'/(D' \cap D)$ is isogenous to $D'D/D$, an algebraic subgroup of the linear group G/D, so $D'/(D' \cap D)$ is linear. Since $D \subset C$ (Theorem 13, Corollary 1), $D' \cap D$ is a normal subgroup of G. If τ is the natural homomorphism from G to $G/(D' \cap D)$ we have $\tau G = (\tau L)(\tau D')$. Since τL and $\tau D'$ are linear, so is τG. Hence $D' \cap D \supset D$; that is $D' \supset D$. If H is the kernel of a rational homomorphism from D to a linear group, then D/H is linear. But $(G/H)/(D/H) = G/D$ is linear, so G/H is linear, giving $H \supset D$; i.e. $H = D$. Finally, let n be an integer > 0 and let σ be the rational homomorphism of D into itself given by $\sigma x = x^n$. Then each element of $D/\sigma D$ has an order dividing n, so any rational homomorphism from this group to an abelian variety is trivial. Hence $D/\sigma D$ is linear. Thus $\sigma D = D$, so σ has a finite kernel.

It should be remarked that D is not necessarily an abelian variety. For example, let the universal domain have characteristic zero and let G be the generalized jacobian variety of an ellipitic curve with an ordinary cusp. Then G is commutative, contains an algebraic subgroup H which is biregularly isomorphic to G_a, and G/H is the jacobian variety of the curve. If Γ is a connected algebraic subgroup of G that does not contain H, then $\Gamma \cap H = \{e\}$. Hence if $\dim \Gamma > 0$, the homomorphism from G to G/H induces a biregular isomorphism between Γ and G/H, contradicting the nonexistence of a regular cross section for the map $G \to G/H$ (cf. [1, Theorem 13]). That is, the only connected algebraic subgroups of G are G, H, and $\{e\}$. Here $D = G$.

COROLLARY 6. *A connected algebraic group is isogenous to the direct product of a connected linear group and an abelian variety if and only if it contains two connected algebraic subgroups of complementary dimensions which are respectively linear and abelian. In this case these subgroups are characterized as the maximal connected linear algebraic subgroup and the maximal abelian subvariety respectively of the given group. Any connected algebraic subgroup and any rational homomorphic image of a group of this type is also of this type.*

Any connected algebraic group which is isogenous to such a direct product contains algebraic subgroups which are isogenous to the direct factors, proving half of the first statement. Conversely, if the connected algebraic group G contains connected algebraic subgroups L and A, respectively linear and abelian, such that $\dim G = \dim L + \dim A$, then A is contained in the center of G, so the map $\lambda \times \alpha \rightarrow \lambda\alpha$ is a rational homomorphism from $L \times A$ into G. $L \cap A$ being finite, this homomorphism has finite kernel, hence is surjective, so $G = LA$ is isogenous to $L \times A$. L is normal in G. G/L, being isogenous to $A/(A \cap L)$, is abelian, so L is the maximal connected linear algebraic subgroup of G. G/A, being isogenous to $L/(L \cap A)$, is linear, so A contains all abelian subvarieties of G. If τ is any rational homomorphism of $G = LA$, then $\tau G = (\tau L)(\tau A)$, and τL and τA are respectively linear and abelian. Finally, let H be any connected algebraic subgroup of $G = LA$. Then $H/(H \cap A)$ is isogenous to HA/A, an algebraic subgroup of the linear group G/A. Similarly $H/(H \cap L)$ is isogenous to HL/L, an algebraic subgroup of the abelian variety G/L. Thus if we set $A' = (H \cap A)_0$, $L' = (H \cap L)_0$, we get H/A' and H/L' respectively linear and abelian. The common rational homomorphic image $H/L'A'$ of both H/A' and H/L' is therefore both linear and abelian. Thus $H = L'A'$.

Note that not all such groups $G = LA$ are direct products. For example, we can find a connected commutative linear group L, an abelian variety A, and elements $\lambda \varepsilon L$, $\alpha \varepsilon A$ of finite order $n > 1$ and then let G be the factor group of $L \times A$ by the subgroup generated by $\lambda \times \alpha$. Then the images of $L \times 0$ and $e \times A$ in G meet in more than one point.

Corollary 7. *Notations being as in Corollary 5 and H being any normal algebraic subgroup of G, the algebraic group G/H is isogenous to the direct product of a connected linear group and an abelian variety if and only if $H \supset (L \cap D)_0$.*

First suppose that $G/H = \Lambda\Delta$, where Λ is a connected linear group and Δ is an abelian variety. Let L', D' be the inverse images of Λ, Δ respectively in the natural homomorphism from G to G/H. Then $G/L' = (G/H)/(L'/H) = \Lambda\Delta/\Lambda$, which is isogenous to the abelian variety $\Delta/(\Delta \cap \Lambda)$, so G/L' is abelian, whence $L' \supset L$. Similarly, $G/D' = (G/H)/(D'/H) = \Lambda\Delta/\Delta$ is isogenous to the linear group $\Lambda/(\Lambda \cap \Delta)$, so $D' \supset D$. Hence $H \supset (L' \cap D')_0 \supset (L \cap D)_0$. As a result of the previous corollary, to prove the converse it suffices to take $H = L \cap D$. Then $G/H = LD/H = (L/H)(D/H)$. Here L/H is linear and $D/H = D/(L \cap D)$ is isogenous to $DL/L = G/L$ which is abelian, so D/H is an abelian variety. This ends the proof.

PROPOSITION 4. *Let G be a commutative algebraic group defined over the field k and let H be a connected k-closed algebraic subgroup of G. If the characteristic of k is $p \neq 0$ suppose also that H possesses only a finite number of elements of order p. Then H is defined over k.*

Since H is a k-closed variety, it is defined over a purely inseparable algebraic extension k' of k, so we need only consider the case where k has characteristic $p \neq 0$. Then the rational endomorphism of H given by $h \rightarrow p \cdot h$ has finite kernel, hence is surjective, so $h \rightarrow p^{\nu} \cdot h$ is surjective for any $\nu \geqq 0$. We now use an idea of Chow. If x is generic for H over k', so is $p^{\nu} \cdot x$. Fix an integer μ so large that $k(k(x))^{p^{\mu}}$ is separably generated over k and then take ν so large that the zero-cycle $p^{\nu}(x)$ on G is rational over $k(k(x))^{p^{\mu}}$. By [6, Theorem 1, p. 15] the point $p^{\nu} \cdot x$ is rational over $k(k(x))^{p^{\mu}}$, so $k(p^{\nu} \cdot x)$ is separably generated over k. Also, k' is algebraically closed in $k'(x) \supset k(p^{\nu} \cdot x)$, so k is algebraically closed in $k(p^{\nu} \cdot x)$. Therefore $k(p^{\nu} \cdot x)$ is a regular extension of k. Hence H is defined over k.

COROLLARY. *Let the connected algebraic group G be defined over k. Then $(L \cap D)_0$ and the maximal abelian subvariety of G are also defined over k.*

Each of these subgroups is left invariant by any k-automorphism of the universal domain, hence is k-closed. But each is an algebraic subgroup of the commutative group D, which is defined over k and has only a finite number of elements of any given finite order.

NORTHWESTERN UNIVERSITY.

REFERENCES.

[1] M. Rosenlicht, "Generalized jacobian varieties," *Annals of Mathematics*, vol. 59 (1954).

[2] ———, "Automorphisms of function fields," *Transactions of the American Mathematical Society*, vol. 79 (1955).

[3] A. Weil, *Foundations of Algebraic Geometry*, American Mathematical Society Colloquium Publications, New York, 1946.

[4] ———, "On algebraic groups of transformations," *American Journal of Mathematics*, vol. 77 (1955).

[5] ———, "On algebraic groups and homogeneous spaces," *American Journal of Mathematics*, vol. 77 (1955).

[6] ———, *Variétés Abéliennes et Courbes Algébriques*, Paris, 1948.

[7] O. Zariski, "Foundations of a general theory of birational correspondences," *Transactions of the American Mathematical Society*, vol. 53 (1943).

ON THE ARTIN ROOT NUMBER.* [1]

By B. Dwork.

Let X be an arbitrary character of the Galois group, $G(K/k)$, of a normal extension, K, of an algebraic number field k. The possibility of determining the arithmetic structure of the Artin root number, $W(X)$, defined by the functional equation [1] of the Artin L-series, $L(s, X, K/k)$, is suggested by the recent arithmetic characterization of $W(X)$ for X linear [2, 3, 4]. If X is linear then $W(X)$ shall be referred to as an "abelian root number."

It is the purpose of this paper to show that much of the abelian theory goes over with little modification. The main result is that if $\mathfrak{p}$ is a prime of k and if $X_\mathfrak{p}$ is the character of the local [2] Galois group, $G(Kk_\mathfrak{p}/k_\mathfrak{p})$, obtained from the restriction of X to a $\mathfrak{p}$ decomposition subgroup of $G(K/k)$ by means of the natural isomorphism between the local Galois group and the decomposition subgroup, then to within a multiplicative factor, ± 1, there exists a well defined local root number, $W(X_p)$, with factor group, linearity and induced character properties such that $W(X) = \prod W(X_\mathfrak{p})$, the product being over all primes of k. Thus (except for the question of sign to be discussed later) it is enough to determine the arithmetic structure of "irreducible local root numbers," $W(\theta)$, where θ is an irreducible character of $G(Kk_\mathfrak{p}/k_\mathfrak{p})$. If $\mathfrak{p}$ is a finite prime and $f(\theta)$ is the (local) conductor of θ, let $m(\theta) = (\mathrm{ord}_\mathfrak{p} f(\theta))/\theta(1)$. It will be shown that $W(\theta)$ is a root of unity unless $m(\theta) = 1$ in which case it is the ratio between a classical Gauss sum and its absolute value.

An immediate consequence is the integrality of Galois Gauss sums as conjectured by Hasse [5], p. 40. Hasse's conjecture (op. cit.) concerning the field in which $W(X)$ lies is treated in Theorem 7 and its corollaries.

Some remarks about the presentation are in order. While the group

* Received November 3, 1955.

[1] A summary of most of these results appears in the author's "The local structure of the Artin root number," *Proceedings of the National Academy of Science, U. S. A.*, vol. 41 (1955), pp. 754-756. Theorem 4a and its consequences form the substance of the author's dissertation, "On the root number in the functional equation of the Artin-Weil L-series, Columbia University (1954), (unpublished).

[2] "Global" and "local" are used to distinguish between algebraic number fields and their completion under a valuation (Archimedean or non-Archimedean).

theoretical discussion of Section 1 may for our immediate purpose be restricted to nilpotent groups, the more extended result has a bearing on the problem of obtaining the group theoretical properties of the Artin root number without analysis. It should also be noted that the proof of Theorem 4 below is reserved for a future paper so as to avoid excessive preoccupation at this time with purely arithmetic computations.

I am indebted to John Tate for his advice and encouragement during this investigation.

1. Group theoretical considerations. Artin's concept of extending functions defined on linear characters is easily abstracted (cf. [5], § 1).

Definition 1. A set of finite groups is said to be a *family* if it is closed under the process of forming subgroups and factor groups, it being understood that if G is a group in the set and L is a subgroup of G which contains an invariant subgroup, H, of G then L/H is identified with the image of L under the natural mapping of G into G/H, the identification to be done by means of the natural isomorphism between the two groups.

In the statement of the above definition, if L is a subgroup of G which does not contain H then $L/(H \cap L)$ is not to be identified with the image of L in G/H. In discussing families our main concern is with the characters (i.e. the traces of matrix representations with coefficients in the field of complex numbers) of the groups in the family and in particular with the mappings of the characters into some fixed abelian group. The purpose of the identification in the definition is simply to impose a natural restriction on these mappings. Of course this is achieved only if the sets of characters of identified groups are also identified in the obvious way and this further identification is to be understood. For the purpose of this section there is no need to identify the naturally isomorphic groups, G/H and $(G/N)/(H/N)$, where G is a group in the family and H is an invariant subgroup of G which contains N, an invariant subgroup of G.

Definition 2. A function, F_0, defined on the set of linear characters of the groups in a family, Δ, and taking its values in some fixed, abelian, (multiplicative) group is said to be *extendable with respect to* Δ (or Δ-extendable) if it can be extended to a function, F, on the set of all characters of groups in Δ with linearity, factor group and induced character properties. Specifically, if $G \varepsilon \Delta$ and L is a subgroup of G then

(a) If L is invariant in G and θ is a character of G/L then $F(\theta)$

$=F(\theta \circ \phi)$, $\theta \circ \phi$ being the character of G obtained by composing θ with ϕ, the natural homomorphism of G onto G/L.

(b) If X and X' are characters of G then $F(X+X')=F(X)F(X')$.

(c) If θ is a character of L and X is the character of G induced by θ then $F(\theta)=F(X)$.

It may be noted that condition (c) implies

(c′) If X is a character of L and $\sigma \varepsilon G$ then $F(X^\sigma)=F(X)$, where X^σ denotes the character $x \rightarrow X(\sigma x \sigma^{-1})$ of the group $\sigma^{-1}L\sigma$.

It is an immediate consequence of Brauer's fundamental theorem on induced characters, [6], that if F exists then it is completely determined by F_0.

While the theory of non-abelian L-series, [1], depends entirely upon the concept of extendable functions, the purely group theoretical problem of characterizing such functions has received no attention. The solution of this problem for families of solvable groups or at least for groups satisfying the conditions imposed by Hilbert theory on the galois groups in local number theory would give the group theoretical structure of the Artin root number by purely arithmetic methods. Unfortunately we can at this time give the solution only for supersolvable groups (i. e. groups whose principal series have cyclic prime factor groups). Some elementary properties of solvable groups are needed before the result can be stated.

LEMMA 1. *If G is a finite solvable group, H a subgroup of prime index, p, H' the maximal subgroup of H which is invariant in G then*

(a) *There exists a unique subgroup G' of G which contains H' as a subgroup of index p.*

(b) *H/H' and G/G' are cyclic groups of equal order which divides $p-1$.*

(c) *There exists an element, x, of G such that $x \notin H$, $x^p \varepsilon H$.*

(d) *The $(p-1)$ non-trivial linear characters of G' which are trivial on H' are permuted by the inner automorphisms of G so that the domains of transitivity contain $m=(H:H')$ elements.*

Proof. It may be assumed that H is not a normal subgroup of G, hence H is its own normalizer in G. It follows that H has exactly p conjugates, $H=H_1, H_2, \cdots, H_p$. For $x \varepsilon G$ let T_x be the permutation, $H_i \rightarrow xH_ix^{-1}$ of the conjugates of H. $x \rightarrow T_x$ is a representation of G onto a transitive permutation group on p elements. The kernel of the representation consists of all

$x \varepsilon G$ such that $xyHy^{-1}x^{-1} = yHy^{-1}$ for all $y \varepsilon G$. As H is its own normalizer, the kernel is the intersection of the conjugates of H, i. e.: H'. Statements (a) and (b) now follow from the well known properties of transitive solvable groups of permutations on a prime number of elements [7], p. 77. Let x be any element of G' not in H'. Statement (c) follows from $G' \cap H = H'$. For (d) let Y be a non-trivial linear character of G' which is trivial on H'. As it is clear that Y has at most m conjugates in G, it is enough to show them to be distinct, i. e. if $x \varepsilon G$, $Y(xyx^{-1}) = Y(y)$ for all $y \varepsilon G'$ then $x \varepsilon G'$. But this hypothesis implies that $\bar{x} = x \bmod H'$ commutes with each element of G'/H' so that G'/H' lies in the center of the group, $\bar{G}$, generated by it and $\bar{x}$, whence $\bar{G}$ is abelian as $\bar{G}/(G'/H')$ is cyclic. It follows from the previously mentioned theory of permutation groups on a prime number of elements that $\bar{G} = G'/H'$, which proves the assertion.

Definition 3. A set of groups (G, H, G', H') written in this order shall be referred to as a $(p, p-1)$ configuration if the groups are related in the manner indicated in the lemma.

Definition 4. A set of groups (G, G_1, G_2, G_0) shall be referred to as a (p, p) configuration if p is prime, G_0 is an invariant subgroup of G, G/G_0 is an abelian (p, p) group and G_1 and G_2 are distinct subgroups of G which contain G_0 as a subgroup of index p.

A family, Δ, is said to be supersolvable if every group in the family is supersolvable. The main result of this section is that the (p, p) and $(p, p-1)$ configurations are the basic units from which the relations between the characters of the groups in a supersolvable family may be determined; specifically:

THEOREM 1. *If Δ is a supersolvable family and if F is a function defined on the set of linear characters of the groups in Δ having the following properties:*

(a) *F is invariant under the transformations of the linear characters produced by inner automorphisms of the groups in Δ* (cf. Definition 2, (c')).

(b) *If (G, G_1, G_2, H) is a (p, p) configuration of groups in Δ, H is abelian and θ_i $(i = 1, 2)$ is a linear character of G_i which induces a given irreducible character of G then $F(\theta_1) = F(\theta_2)$.*

(c) *If (G, G_0, G_1, H) is a $(p, p-1)$ configuration of groups in Δ, H is abelian, θ is a linear character of G, $\theta_0 = \theta | G_0$,[3] $\theta_1 = \theta | G_1$ and $Y_1, \cdots, Y_r$*

[3] Read: The restriction of θ to G_0.

is a minimal set of non-trivial linear characters of G_1, *trivial on* H, *whose* G *conjugates cover the set of all such characters* (cf. Lemma 1), *then* $F(\theta_0) = F(\theta) \prod_{i=1}^{r} F(\theta_1 Y_i)$.

(d) F *has the factor group property* (cf. Definition 2(a)) *for linear characters.*

Then F *is* Δ*-extendable. Regardless of* Δ *these conditions are necessary for extendability.*

This theorem is a direct consequence of Theorems 1A and 1B below. In the discussion of these theorems it is to be understood that F and Δ satisfy the conditions of Theorem 1.

THEOREM 1A. *If* X *is an irreducible character of* $G \varepsilon \Delta$ *which is induced by a linear character* θ_i *of a subgroup* G_i *of* G $(i = 1, 2)$ *then*

$$F(\theta_1) = F(\theta_2). \tag{1}$$

Proof. Let H be a maximal abelian subgroup of $G_1 \cap G_2$ which is invariant in G. The theorem is trivial if $(G:H) = 1$. The proof is by induction on the index $(G:H)$, the idea of the proof being to use the induction hypothesis to replace the given group, G, by one in which hypotheses (b) and (c) may be applied. As θ_1 and θ_2 induce the same irreducible character of G, it follows from Mackey, [8], that there exists $x \varepsilon G$ such that θ_2 and $\theta_1{}^x$ (exponentiation as in Definition 2(c′)) coincide on $G_2 \cap x^{-1}G_1x$. Hence by hypothesis (a) it may be assumed that θ_1 and θ_2 coincide on $G_0 = G_1 \cap G_2 \supset H$. As θ_1 and θ_2 induce irreducible characters of G_1G_2, it follows from the converse part of the previously mentioned theorem of Mackey that they induce the same character of G_1G_2. Hence by the induction hypothesis it may be assumed that $G = G_1G_2$. Let δ be the common restriction of θ_1 and θ_2 to H. δ is invariant under G_1 and G_2 and therefore under G. As G is supersolvable there exists an invariant subgroup, M, of G which contains H as a subgroup of prime index, p. Let Φ be a matrix representation of G whose character is X, then $\Phi | H = \delta I_d$, where I_d is the unit matrix of rank $d = X(1)$. As δ is invariant under G, $\Phi(H)$ lies in the center of $\Phi(G)$, hence certainly in the center of $\Phi(M)$ while $\Phi(M)/\Phi(H)$ is cyclic, whence $\Phi(M)$ is abelian. As $\Phi(G)/\Phi(M)$ is supersolvable it follows from Taketa, [9], that the character, $A \rightarrow \operatorname{Trace} A$, of the group $\Phi(G)$ is induced by a linear character of a subgroup of $\Phi(G)$ which contains $\Phi(M)$. It follows without difficulty that X is induced by a linear character of a subgroup of G which contains M. Without loss in generality it may be assumed that G_2

contains M. It may be further assumed that M is not a subgroup of G_1 as otherwise by hypothesis (d) and the identifications of Definition 1, θ_1 and θ_2 may be replaced by linear characters of the subgroups G_1/M^c, G_2/M^c of the factor group, G/M^c, ($M^c =$ commutator subgroup of M) which induce the same irreducible character of the factor group, whence (1) follows from the induction hypothesis as M/M^c is an abelian subgroup of $(G_1/M^c) \cap (G_2/M^c)$ which is invariant in G/M^c and of index $(G:M) < (G:H)$. (The identifications of Definition 1 are used only in this argument. Reference is made to this argument at several points in the remainder of the proof.)

Hence $G_1 \cap M = H$ so that G_1 is a subgroup of G_1M of index p. Let X' be the character of G_1M induced by θ_1. X' is irreducible and its restriction to H is $p\delta$. As δ is certainly invariant under G_1M it follows from a previous argument that X' is induced by a linear character, θ', of a subgroup, G', of G_1M of index p which contains M. If $G \neq G_1M$ then by the induction hypothesis $F(\theta_1) = F(\theta')$ while a previous argument shows that $F(\theta') = F(\theta_2)$ as $G' \cap G_2 \supset M$.

Hence it may be assumed that $G = G_1M$, $G_2 \supset M$, $p = (G:G_1) = (G:G_2)$, $G_1 \not\supset M$ and θ_1 and θ_2 coincide on $G_0 = G_1 \cap G_2$. Clearly $G_0 \cap M = H$ so that $p = (G:G_1) \geqq (G_2:G_0) \geqq (G_0M:G_0) = (M:H) = p$. Hence $(G_1:G_0) = (G_2:G_0) = p$ and $G_2 = G_0M$. If $G_0 = H$ then $(G:H) = p^2$, whence G/H is abelian but not cyclic (as otherwise $G_1 = G_2 = G_1G_2 = G$) so that (G, G_1, G_2, H) is a (p,p) configuration so that (1) follows from hypothesis (b). Thus it may be assumed that $G_0 \neq H$ and by the argument used to justify the assumption that $G_1 \not\supset M$, it may be further assumed that no subgroup of G_0 properly containing H is invariant in G.[4]

Summarizing: It is enough to prove (1) for the case in which $G = G_1G_2$, $p = (G:G_1) = (G:G_2) = (G_1:G_0) = (G_2:G_0)$, $G_0 = G_1 \cap G_2 \supset H$ (strict inclusion), H an abelian invariant subgroup of G, $\theta_1 | G_0 = \theta_2 | G_0$, $p^2 | (G:H)$, no subgroup of G_0 properly containing H is invariant in G.

As no further use shall be made of the group M, there will be no need in the remainder of the proof to repeat for G_1 arguments applicable to G_2 and conversely. To complete the proof it is necessary to examine the structure of G more closely. We assert: If U is a subgroup of G_2 which is invariant in G then $U' = U \cap G_0$ is also invariant in G. To prove this let $s_1, \cdots, s_p$ be a set of representatives of the right cosets of G_0 in G_2 and let $t_1, \cdots, t_p$ be a set of representatives of the right cosets of G_0 in G_1. If $u \in U'$ then $t_i u t_i^{-1} \in G_0$, whence $X(u) = \sum_{i=1}^{p} \theta_2(t_i u t_i^{-1}) = p\theta_1(u)$. Hence $p\theta_1(u) = \sum_{i=1}^{p} \theta_1(s_i u s_i^{-1})$.

[4] This completes the proof if G is nilpotent.

If $s_i u s_i^{-1} \varepsilon G_1$ then $\theta_1(s_i u s_i^{-1}) = \theta_1(u)$, whence $s_i u s_i^{-1} \varepsilon G_1$ for $i = 1, \cdots, p$ and therefore $s_i U' s_i^{-1} \subset G_1$. It follows that $s_i U' s_i^{-1} = U'$ for each i and as U' is invariant in G_1 the assertion follows.

Let Z_i $(i = 1, 2)$ be the maximal subgroup of G_i which is invariant in G. It follows from the above assertion that $Z_i \cap G_0$ is invariant in G and contains H and therefore is H. From Lemma 1, $(G:Z_i)$ divides $p(p-1)$ but p^2 divides $(G:H)$, hence p divides $(Z_i:H)$. As $(G_1Z_2:Z_2) = (G_1:H)$ it follows that $(G:H) \geqq (G_1Z_2:H) = (G_1Z_2:Z_2)(Z_2:H) = (G_1:H)(Z_2:H) \geqq p(G_1:H) = (G:H)$. Hence $G_1Z_2 = G$ and $(Z_2:H) = p$. Likewise $(Z_1:H) = p$ and $G_2Z_1 = G$. It now follows that (G_1, G_0, Z_1, H) is a $(p, p-1)$ configuration, for if M is a subgroup of G_0 which is invariant in G_1 then MZ_2 is invariant under conjugation by elements of G_1 and by elements of Z_2 and therefore by elements of $G_1Z_2 = G$, whence $MZ_2 = Z_2$ so that $M \subset G_0 \cap Z_2 = H$ and therefore H is the maximal such subgroup of G_0. Likewise (G_2, G_0, Z_2, H) is a $(p, p-1)$ configuration. Furthermore $H \subset Z_1 \cap Z_2 \subset Z_1 \cap G_0 = H$ so that letting $Z = Z_1Z_2$, we have $(Z:H) = p^2$ and therefore (Z, Z_1, Z_2, H) is a (p, p) configuration.

Let $\phi_1 = \theta_1 | Z_1$ and let s be an element of Z_2 which generates the factor group Z_2/H. Set $L = G_1 \cap s^{-1}G_1 s$, then L and sLs^{-1} are subgroups of G_1 which contain Z_1 and are of the same index in G_1. As G_1/Z_1 is cyclic it follows that they are equal, whence L is an invariant subgroup of G, i. e. $L = Z_1$. As θ_1 induces an irreducible character of G it follows from Mackey (op. cit.) that θ_1 and θ_1^s do not coincide on Z_1 but do coincide on H as δ is invariant under G. As ϕ_1 is invariant under G_1 it follows that ϕ_1 has exactly p conjugates, $\phi_1 Y_0, \phi_1 Y_1, \cdots, \phi_1 Y_{p-1}$, where $Y_0, \cdots, Y_{p-1}$ are the linear characters of Z_1 which are trivial on H. A similar statement holds for $\phi_2 = \theta_2 | Z_2$. It is easily verified that ϕ_1 induces the character $X | Z$ of Z. $X | Z$ is irreducible as now follows from the converse part of the last mentioned theorem of Mackey (it is for this that s is chosen in Z_2). It follows from hypothesis (b) that

$$F(\phi_1) = F(\phi_2). \tag{2}$$

Furthermore, letting $\theta_0 = \theta_1 | G_0$, it follows from hypothesis (c) that

$$F(\theta_0) = F(\theta_1) \prod_{i=1}^{r} F(\phi_1 Y_i) \tag{3}$$

for a suitable indexing of the Y_i. From Lemma 1, $r = (p-1)/(G_0:H)$, but the $\phi_1 Y_i$ are conjugates of ϕ_1, whence by hypothesis (a),

$$F(\theta_0) = F(\theta_1)[F(\phi_1)]^r. \tag{4}$$

Clearly (4) remains valid if the subscript 1 is replaced by 2, hence (1) follows from (2).

This completes the proof of Theorem 1A. F may be extended in a natural way to all irreducible characters of groups in Δ as if X is such a character then it is induced by a linear character θ of a subgroup and while the choice of θ need not be unique, $F(\theta)$ depends only upon X. The extension to composite characters so that the linearity condition is satisfied is obvious. The same symbol, F, will be used to denote the extended function. The factor group property follows easily from hypothesis (d). Some preliminary results are needed to verify the induced character property (and so complete the proof of Theorem 1). In these elementary lemmas the groups referred to are understood to be arbitrary finite groups.

LEMMA 2. *Let X be an irreducible character of a group G whose restriction to an invariant subgroup, H contains a linear character, δ. If G' is the subgroup of G which leaves δ invariant then $X|G'$ contains just one irreducible character, Θ, whose restriction to H contains δ. Furthermore X is induced by Θ.*

Proof. Certainly $X|G'$ contains an irreducible character, Θ, of G' whose restriction to H contains δ. Clearly $\Theta|H=\Theta(1)\delta$. If $t \varepsilon G$, $t \not\varepsilon G'$ then $\delta^t \neq \delta$ so that $\Theta^t|H$ and $\Theta|H$ have in common no irreducible character of H, hence certainly the restrictions of Θ^t and Θ to $G' \cap t^{-1}G't$ have no irreducible character of that group in common. It follows from the previously mentioned results of Mackey that Θ induces X. Hence $X|H$ contains δ exactly $\Theta(1)$ times so that no other irreducible character of G' which occurs in $X|G'$ can have δ in its restriction to H.

LEMMA 3. *Let X be the character of a group G which is induced by a linear character, θ, of a subgroup L. If δ is the restriction of θ to an invariant subgroup, H, of G contained by L and if G' is the subgroup of G which leaves δ invariant then the characters, Θ of G' and X of G, induced by θ, have decompositions, $\Theta=\sum_{i=1}^{r} a_i\Theta_i$, $(a_i \neq 0)$ $X=\sum_{i=1}^{r} a_iX_i$, X_i is induced by Θ_i, into distinct irreducible characters of their respective groups.*

Proof. Clearly L is a subgroup of G'. Let $\Theta=\sum_{i=1}^{r} a_i\Theta_i$ be the decomposition of Θ into distinct irreducible characters of G': As Θ induces X, each irreducible character, X_j, of G occurring in X lies in the character of G induced by one of the Θ_i. If Θ_i and X_j are so related then by Frobenius,

Θ_i occurs in $X_j | G'$, whence by Lemma 2, Θ_i is the only irreducible character of G' which can be so related and furthermore Θ_i induces X_j. The lemma follows directly.

The induced character property may now be demonstrated.

THEOREM 1B. *Let θ be a character of a subgroup, L, of a group $G \varepsilon \Delta$. Let X be the character of G induced by θ, then $F(X) = F(\theta)$.*

Proof. It may be assumed that θ is irreducible and therefore it may be assumed that θ is linear. Let H be a maximal abelian subgroup of L which is invariant in G. The theorem is trivial if $(G:H) = 1$ and the proof is by induction on this index. Once again the idea of the proof is to use the induction hypothesis to replace G by a group in which hypothesis (c) may be used. Let $\delta = \theta | H$ and let G' be the subgroup of G which leaves δ invariant. Using the notation and result of Lemma 3 it is clear that $F(X) = F(\Theta)$, whence by the induction hypothesis it may be assumed that $G' = G$, i. e. δ is invariant under G. As before it follows from the induction hypothesis and the factor group property that it may be assumed that H is the maximal invariant subgroup of G in L. As before there exists an invariant subgroup, M, of G which contains H as a subgroup of prime index, p. From the preceding remark $L \not\supseteq M$, whence $L \cap M = H$. Let $LM = N$, then $(N:L) = p$. Let X' be the character of N which is induced by θ. If $N \neq G$ then by the induction hypothesis $F(\theta) = F(X')$. As $X' | H = X'(1)\delta$ it follows from a previous argument that each irreducible character of N in X' is induced by a linear character of a subgroup of N which contains M. It follows from the linearity and factor group properties of F and the induction hypothesis that $F(X') = F(X)$. Hence it may be assumed that $G = ML$, $(G:L) = p$. As H is the maximal invariant subgroup of G in L it follows that (G, L, M, H) is a $(p, p-1)$ configuration. Let $m = (L:H) = (G:M)$ then $m | (p-1)$. We assert that there exists a linear character X_0 of G whose restriction to H is δ. This is clear if $m = 1$ as then G/H is cyclic, whence from the invariance of δ under G it follows that δ is trivial on the commutator subgroup of G. On the other hand if $m > 1$ then $sLs^{-1} \subset L$ for no element $s \varepsilon M$, $s \not\varepsilon H$. Let t be an element of L whose coset mod H generates the cyclic group, L/H. Then $X(t) = \theta(t) \neq 0$. But each irreducible character, X_i, of G occurring in X is induced by a linear character of a subgroup, G_i, which contains M. As G/M is cyclic, G_i is invariant in G, whence $X_i(t) = 0$ unless $t \varepsilon G_i$. Hence there exists i such that $t \varepsilon G_i$ so that $G_i = G$ and therefore $X_i(1) = 1$. Having shown that X_0 exists it follows from the reciprocity law that $\theta = X_0 | L$. Let $\theta_0 = X_0 | M$, then θ_0 is invariant

under G. As $X|M$ is the character of M induced by δ, it follows easily that $X|M = \theta_0 + \theta_0 Y_1 + \cdots + \theta_0 Y_{p-1}$, where $Y_1, \cdots, Y_{p-1}$ are the nontrivial linear characters of M which are trivial on H. By Lemma 1, Y_i has m distinct conjugates in G, whence the same holds for $\theta_0 Y_i$ and therefore by Mackey, $\theta_0 Y_i$ induces an irreducible character of G of degree m. Each of the $r = (p-1)/m$ distinct irreducible characters of G of this type must occur at least once in X and therefore $X = X_0 + \sum_{i=1}^{r} X_i$, X_i induced by $\theta_0 Y_i$; $Y_1, \cdots, Y_r$ chosen so that their G conjugates cover the full set $Y_1, \cdots, Y_{p-1}$. By definition $F(X) = F(X_0) \prod_{i=1}^{r} F(\theta_0 Y_i)$, but the right side is $F(\theta)$ (hypothesis (c)). This completes the proof of Theorem 1B and therefore of Theorem 1, the necessity of the conditions being now clear.

The solvable group of lowest order which is not supersolvable is the tetrahedron group of order 12. Supersolvable groups are characterized by the property: Every maximal subgroup is of prime index [13]. For this reason Lemma 1 is adequate only for the study of supersolvable families. The extension of the theorem to solvable groups requires an examination of the situation discussed in Lemma 1 with H a subgroup of G of index p^r (p prime). This can be done if enough is known about solvable groups of permutations of p^r elements.

2. Field theoretic considerations. Theorem 1 permits the construction of local root numbers of characters of supersolvable local Galois groups. To remove the restriction of supersolvability without a deeper group theoretical investigation it is necessary to make use of the analytically derived group theoretical structure of the (global) root number as explained by Theorem 3 below. This is done by means of the relations between local and global number fields, specifically the existence of global number fields whose local completions may to some extent be preassigned.

THEOREM 2. *Let A be a finite normal non-cyclic extension of a local number field, B. If m (>1) is the degree over B of an intermediate field then there exists an algebraic number field, k, with normal overfield, K, such that the set, S, of all primes of k which have just one prime divisor in K has either one or m elements and in either case B is (topologically[5]) isomorphic to the completion of k at each prime in S and A is isomorphic to the completion of K at each prime of K which divides a prime in S. Furthermore the fields K, k may be so chosen that for each prime, $\mathfrak{P}$, of K*

[5] The topology is given by the valuation.

which divides a prime in S there exists a $\mathfrak{P}$-topological isomorphism $\Phi_{\mathfrak{P}}$ of K into A which maps k into B such that if $\Sigma \to \Sigma_{\mathfrak{P}}$ is the isomorphism of $G(K/k)$ onto $G(A/B)$ defined by

$$(\Sigma_{\mathfrak{P}} \circ \Phi_{\mathfrak{P}})(x) = (\Phi_{\mathfrak{P}} \circ \Sigma)(x) \text{ for all } x \varepsilon K$$

then the mapping $\Sigma \to \Sigma_{\mathfrak{P}}$ remains unchanged as $\mathfrak{P}$ runs through the primes of K which divide a prime in S.

Proof. Let n be the degree of A over B. Certainly there exists an algebraic number field, k, with normal extension, K, of degree n such that B is isomorphic to the completion of k at some prime, $\mathfrak{p}_0$, of k and such that A is isomorphic to the completion of K at the prime of that field which divides $\mathfrak{p}_0$. Let S_0 be the set of all primes of k other than $\mathfrak{p}_0$ which have just one prime divisor in K. S_0 is finite as K is a non-cyclic extension of k. It may be assumed that S_0 is not empty. By hypothesis there exists a subfield, L, of K such that $m = \text{degree } L/k > 1$. Let v be an integer of L which generates L over k. Let f be the irreducible monic polynomial with coefficients in k which is satisfied by v. Let $a_1, \cdots, a_m$ be any set of distinct integers in k and let $\prod_{i=1}^{m} (x - a_i) = f_0(x)$. By the approximation theorem [10, page 8] there exists a polynomial, g, of degree m with coefficients in k which is so close to f at the primes of S_0 and to f_0 at $\mathfrak{p}_0$ that

(a) g splits in $k_{\mathfrak{p}_0}$, the $\mathfrak{p}_0$ completion of k.

(b) If $\mathfrak{p} \varepsilon S_0$ and if $v_1, \cdots, v_m$ are the roots of f and $b_1, \cdots, b_m$ are the roots of g in a field containing $k_{\mathfrak{p}}$, then for suitable choice of indices, $k_{\mathfrak{p}}(v_i) = k_{\mathfrak{p}}(b_i)$, $1 \leqq i \leqq m$.

Let w be a root of the polynomial, g, in an extension field of K. Let $F = k(w)$ and let $E = KF$. We assert that the fields F, E have the required properties. For $\mathfrak{p} \varepsilon S_0$, f is irreducible in $k_{\mathfrak{p}}$ and therefore by (b), g is irreducible in $k_{\mathfrak{p}}$, hence g is irreducible in k. Thus degree $F/k = m$ and $\mathfrak{p}$ has just one prime divisor in F. By (a), $k_{\mathfrak{p}_0} \supset F$ so that $\mathfrak{p}_0$ has m prime divisors, $\mathfrak{q}_1, \cdots, \mathfrak{q}_m$ in F and for each of them $F_{\mathfrak{q}_i} = k_{\mathfrak{p}_0} \cong B$ while $EF_{\mathfrak{q}_i} = Kk_{\mathfrak{p}_0} \cong A$, $(1 \leqq i \leqq m)$. Hence

$$\text{degree } E/F \geqq \text{degree } A/B = n = \text{degree } K/k \geqq \text{degree } E/F.$$

It follows that $n = \text{degree } E/F$ and that each of these primes of F has just one prime divisor in E. Thus it suffices to show that the remaining primes of F have more than one prime divisor in E. If $\mathfrak{q}$ is one of these remaining

primes of F then $\mathfrak{q}$ divides a prime, $\mathfrak{p}$, of k which either lies in S_0 or has more than one prime divisor in K.

Case 1. $\mathfrak{p} \varepsilon S_0$. As previously noted $\mathfrak{q}$ is the only prime divisor of $\mathfrak{p}$ in F. Since f splits in K and therefore in $Kk_{\mathfrak{p}}$, it follows that g splits in $Kk_{\mathfrak{p}}$ so that $Kk_{\mathfrak{p}} \supset F_{\mathfrak{q}}$. Hence $EF_{\mathfrak{q}} = KF_{\mathfrak{q}} = Kk_{\mathfrak{p}}$ is of degree n over $k_{\mathfrak{p}}$ and therefore of degree n/m over $F_{\mathfrak{q}}$ which shows that $\mathfrak{q}$ has m prime divisors in E.

Case 2. $\mathfrak{p} \not\varepsilon S_0$. As $\mathfrak{p}$ has more than one prime divisor in K,

$$n > \operatorname{degree} Kk_{\mathfrak{p}}/k_{\mathfrak{p}} \geqq \operatorname{degree} EF_{\mathfrak{q}}/F_{\mathfrak{q}}$$

(as $EF_{\mathfrak{q}}$ is the composition of $Kk_{\mathfrak{p}}$ with $F_{\mathfrak{q}}$), which shows that $\mathfrak{q}$ has more than one prime divisor in E.

The last assertion of the theorem follows from the fact that for $1 \leqq i \leqq m$, the prime $\mathfrak{Q}_i$ of E which divides the prime q_i of F also divides that prime, $\mathfrak{P}_0$, of K which divides the prime, $\mathfrak{p}_0$, of k. Furthermore the local degree at $\mathfrak{Q}_i$ of E over K is 1. Let ϕ be a $\mathfrak{P}_0$-topological isomorphism of K into A which maps k into B and let $\Sigma \to \bar{\Sigma}$ be the isomorphism of $G(E/F)$ onto $G(A/B)$ defined by

$$(\bar{\Sigma} \circ \phi)(x) = (\phi \circ \Sigma)(x) \text{ for all } x \varepsilon K.$$

It is clear that ϕ can be extended to Φ_i, a $\mathfrak{Q}_i$-topological isomorphism of E into A. Φ_i maps F into B and the mapping $\Sigma \to \Sigma_i$ of $G(E/F)$ onto $G(A/B)$ defined by means of Φ_i is the same as the mapping $\Sigma \to \bar{\Sigma}$. This completes the proof of the theorem.

The basic problem of this section is that of characterizing functions defined on linear characters of local Galois groups which can be extended with respect to the family of all local Galois groups. Theorem 3 below (together with Theorem 1) gives a partial solution of this problem which is adequate for the theory of root numbers. Before stating the result we pause to explain the terminology.

Let Q be the field of rational numbers and for each prime p of Q (including the infinite one) let Q_p be a p-completion of Q. Let $\mathfrak{T}_p$ (resp.: $\mathfrak{T}$) be the set of all overfields of Q_p (resp.: Q) of finite degree which lie in a fixed algebraic closure of Q_p (resp.: Q), the construction being so performed that no two of these algebraic closures have any element in common. Let $\bar{\mathfrak{T}}$ be the set theoretic union of the disjoint sets, $\mathfrak{T}_p$, where p ranges over all primes of Q. The elements of $\bar{\mathfrak{T}}$ are understood to be topological fields and isomorphisms between two elements will be understood to be topological. $\bar{\mathfrak{T}}$ (resp.: $\mathfrak{T}$) shall be referred to as the *set of all local* (resp.: *global*) *number*

fields. For A, B in $\bar{\mathfrak{T}}$ (resp.: $\mathfrak{T}$), A a normal overfield of B, the Galois group, $G(A/B)$, of A over B is regarded not as an abstract group but rather as being inextricably connected with the pair (A,B) of fields. The usual Galois theoretic identifications concerning subgroups and factor groups being understood, it is clear that the set, $\bar{\mathfrak{G}}$ (resp.: $\mathfrak{G}$), of all such Galois groups is a family in the sense of Definition 1 and shall be referred to as the *family of all local* (resp.: *global*) *Galois groups.* (Precisely as in the purely group theoretical case, $G(CD/D)$ is not to be identified with $G(C/C \cap D)$ as elements of $\bar{\mathfrak{G}}$ (resp.: $\mathfrak{G}$) even when these groups are isomorphic (unless $C \supset D$).)

The set, $\bar{\mathfrak{X}}$ (resp.: $\mathfrak{X}$) of all characters of groups in $\bar{\mathfrak{G}}$ (resp.: $\mathfrak{G}$) shall be referred to as the *set of all characters of local* (resp.: *global*) *Galois groups.* Two characters in $\bar{\mathfrak{X}}$ (resp.: $\mathfrak{X}$) are said to be *equivalent* if one is obtained from the other in the natural way from a field isomorphism. Let Γ be a multiplicative abelian group, fixed throughout the discussion. A mapping, F, of $\bar{\mathfrak{X}}$ (resp.: $\mathfrak{X}$) into Γ which is constant on each class of equivalent characters is said to be a *function defined on all characters of local* (resp.: *global*) *Galois groups.* A similar interpretation is to be given to the *set of all linear characters of local* (resp.: *global*) *Galois groups* and to the concept of a *function defined on all linear characters of local* (resp.: *global*) *Galois groups.* A function defined on all linear characters of local (resp.: global) Galois groups will be said to be extendable with respect to a given family of local (resp.: global) Galois groups if the restriction of the function to the linear characters of the groups in the family is extendable with respect to the family.

If A is a local number field then a homomorphism of A^* (the multiplicative group of non-zero elements of A) into the unimodular complex numbers wtih kernel of finite index in A^* will be said to be a *multiplicative character of A* (or a *character of A^**). The set of all such homomorphisms as A ranges over $\bar{\mathfrak{T}}$ will be called the *set of all multiplicative characters of local number fields.* Two such characters are said to be *equivalent* if one is transformed into the other by a field isomorphism. A mapping, F, of all multiplicative characters of local number fields into Γ which is constant on each class of equivalent characters is said to be a *function defined on all multiplicative characters of local number fields.* It follows from local class field theory[6] that such functions may be identified with functions defined on all linear

[6] To avoid confusion with the classical norm rest symbol, it is to be understood that a prime element of $k_{\mathfrak{p}}$ (if $\mathfrak{p}$ is a finite prime) is to be associated with automorphisms of abelian over-fields which in the unramified case is simply the Frobenius substitution.

characters of local Galois groups which have the factor group property. This identification is to be understood in the following.

Finally, if K and k are elements of $\mathfrak{T}$, K being normal over k, and if $\mathfrak{p}$ is a prime of k then $G(Kk_{\mathfrak{p}}/k_{\mathfrak{p}})$ will be used to denote an element, $G(A/B)$, of $\bar{\mathfrak{G}}$ such that there exists a $\mathfrak{P}$-topological isomorphism, ϕ, of K onto a dense subset of A which maps k onto a dense subset of B, where $\mathfrak{P}$ is a prime of K which divides $\mathfrak{p}$. If X is a character of $G(K/k)$ then $X_{\mathfrak{p}}$ is to denote the character of $G(A/B)$ which is obtained by transforming (by means of ϕ) the restriction of X to a $\mathfrak{P}$-decomposition subgroup of $G(K/k)$. Of course $X_{\mathfrak{p}}$ is not a uniquely defined element of $\bar{\mathfrak{X}}$ but if F is a function defined on all characters of local Galois groups then $F(X_{\mathfrak{p}})$ is well defined.

Theorem 3. *Let F be a function defined on all linear characters of local Galois groups with the properties:*

(a) *F is extendable with respect to every family of nilpotent local Galois groups.*

(b) *If X is a linear character of a global Galois group, $G(K/k)$, where k is an algebraic number field and K is a (finite) normal overfield, then $F(X_{\mathfrak{p}}) = 1$ for almost all primes, $\mathfrak{p}$, of k ($X_{\mathfrak{p}}$ being defined as indicated above.*

(c) *If in the notation of* (b), *M denotes the function $X \rightarrow \prod_{p} F(X_{\mathfrak{p}})$ defined on the set of linear characters of global Galois groups, the product being over all primes of k, then M is extendable with respect to every family of global Galois groups.*

Then it may be concluded that F is extendable with respect to the family of all local Galois groups.

Proof. It is enough to prove:

If B is a $\mathfrak{p}$-adic number field, A a normal extension of finite degree, and Δ' is the family of groups generated by the Galois group, $G(A/B)$, (i.e. the family of all Galois groups, $G(A'/B')$, where $A \supset A' \supset B' \supset B$, A' normal over B') then F is extendable with respect to Δ'.

The proof is by induction on degree A/B. It may be assumed that the Galois group, $G(A/B)$, is neither cyclic nor a p-group. Let m (> 1) be the degree over B of some intermediate field. Let k and K be the algebraic number fields whose existence has been demonstrated in the last theorem, and let S be the set of all primes of k which have just one prime divisor in K. The number, r, of primes in S is either 1 or m. By the induction

hypothesis, F is extendable with respect to the family of groups generated by the local Galois group, $G(Kk_{\mathfrak{p}}/k_{\mathfrak{p}})$, for each prime $\mathfrak{p}$ of k which is not in S. Let F' denote this extended function for all $\mathfrak{p} \not\in S$. Let Δ be the family of groups generated by $G(K/k)$. If $G_1 \varepsilon \Delta$ then G_1 is the Galois group, $G(U/V)$, of U over V, U and V being fields lying between k and K, U being normal over V. If X is a character of G_1 then for each prime, $\mathfrak{p}$, of k not in S we set

$$H_{\mathfrak{p}}(X) = \prod_{\mathfrak{q}|\mathfrak{p}} F'(X_{\mathfrak{q}}), \tag{4}$$

the product being over all primes, $\mathfrak{q}$, of V which divide $\mathfrak{p}$ and $X_{\mathfrak{q}}$ being related to X in the usual way. For each prime, $\mathfrak{p}$, of k not in S, (4) defines a function on the characters of the groups in Δ. As F' satisfies the conditions of Definition 2 with respect to those local groups of which the $X_{\mathfrak{q}}$ in (4) are characters, it follows from an argument of Artin [11, pages 3-5] that $H_{\mathfrak{p}}$ satisfies the conditions of Definition 2 with respect to the family Δ. If X is a linear character of G_1 then by hypothesis (b),

$$H_{\mathfrak{p}}(X) = 1 \text{ for almost all primes, } \mathfrak{p}, \text{ of } k. \tag{5}$$

By Brauer (loc. cit.), (5) holds for all characters of G_1. Hence if we set

$$H(X) = \prod_{\mathfrak{p} \not\varepsilon S} H_{\mathfrak{p}}(X), \tag{6}$$

the right side being a product over all primes of k not in S, then H is a well defined function of the characters of the groups in Δ which satisfies all the conditions of Definition 2, and therefore (using M to denote the extension of the function, M, (hypothesis (c)))

$$N(X) = M(X)/H(X) \tag{7}$$

defines a function, N, on these characters which also satisfies these conditions, i.e. the restriction of N to the linear characters of the groups in Δ is Δ-extendable.

If X is a linear character of G_1 then

$$N(X) = \prod_{\mathfrak{p} \varepsilon S} \prod_{\mathfrak{q}|\mathfrak{p}} F(X_{\mathfrak{q}}), \tag{8}$$

the combined product being over all primes of V which divide a prime in S. Let $\bar{S}$ be the set of primes of K which divide a prime in S. For each $\mathfrak{P} \varepsilon \bar{S}$ let $\Phi_{\mathfrak{P}}$, a $\mathfrak{P}$-topological isomorphism of K into A, be chosen as in the last assertion of Theorem 2. For each $\Phi_{\mathfrak{P}}$ there exists a one to one correspondence between groups in Δ and groups in Δ' and by the choice of the $\Phi_{\mathfrak{P}}$ this correspondence does not change as $\mathfrak{P}$ runs through $\bar{S}$. Likewise there exists a one to one correspondence $X \leftrightarrow X'$ between the characters of the groups in Δ

and the characters of the groups in Δ' which can be described in terms of $\Phi_{\mathfrak{P}}$ but does not depend upon the choice of $\mathfrak{P}$ in $\bar{S}$. It follows from the topological nature of this construction that if X is the linear character of G_1 in (8) then $F(X') = F(X_{\mathfrak{q}})$ for each prime, $\mathfrak{q}$, of V which divides a prime in S. As each prime in S has just one prime divisor in V, it follows that

(9) $$N(X) = (F(X'))^r \text{ for } X \text{ linear.}$$

It is now clear that the function $F^r: X' \to (F(X'))^r$ on the linear characters of the groups in Δ' is Δ'-extendable. It may be assumed that it is impossible to construct the fields, K, k, so that $r = 1$, hence it may be assumed that $r = m$, the degree over B of a field (not B) lying between A and B. As $G(A/B)$ is not a p-group, the indices of the Sylow subgroups from a set of relatively prime integers each greater than one, which proves the theorem.

While the theorem is adequate for our purposes it is clear that condition (a) is stronger than necessary.

Corollary. *The theorem remains valid if condition* (a) *is replaced by either:*

(a′) *F is extendable with respect to every family of local Galois groups which consists of cyclic groups and p-groups,*

or

(a″) *F is extendable with respect to every family of local Galois groups which consists of p-groups and of the Galois groups of unramified extensions.*

The proof is clear but it should be noted that the last sentence of the proof of the theorem shows why p-groups require special treatment. The special treatment of cyclic groups (or of the Galois groups of unramified extensions) is necessitated by the requirement of finiteness of S_0 in the proof of Theorem 2. It should be observed that (a″) corresponds to an obvious modification of that theorem.

3. Arithmetic considerations. The local abelian root numbers must now be defined.

Definition 5. If θ is a multiplicative character of a local number field, B, then the local abelian root number, $R(\theta)$, is defined by

$$\begin{aligned} R(\theta) &= 1 \quad \text{if } B \text{ is complex,} \\ &= 1 \quad \text{if } B \text{ is real and } \theta(-1) = 1, \\ &= -i \quad \text{if } B \text{ is real and } \theta(-1) = -1, \\ &= (Nf(\theta))^{-\frac{1}{2}} \sum_{x \in U/(1+f(\theta))} \bar{\theta}(x/A_\theta)\phi(x/A_\theta) \quad \text{if } B \text{ is } \mathfrak{p}\text{-adic,} \end{aligned}$$

where (in explanation of the $\mathfrak{p}$-adic case)

$f(\theta)=$ conductor of θ,

$Nf(\theta)=$ absolute norm of $f(\theta)$,

A_θ is an integer of B which generates the ideal, $f(\theta)\mathfrak{D}_B$, $\mathfrak{D}_B$ being the absolute different of B,

ϕ is the "standard" additive character of $B: x \rightarrow \exp(2\pi i Y(S(x)))$, S being the absolute trace mapping B onto Q_p, the corresponding completion of the field of rational numbers, Y being a mapping of Q_p into the rationals such that for each $t \in Q_p$, $Y(t)-t$ is a p-adic integer and $Y(t)$ is a rational number whose denominator is a power of p,

U is the group of units in B,

and the sum is over a set of representatives of the cosets in U of the subgroup of all units congruent to 1 modulo $f(\theta)$, it being understood that if θ is unramified then $1+f(\theta)=U$.

It follows from the arithmetic characterization of abelian root numbers, [4], that $R(\theta)$ is a unimodular complex number in any case and is even a root of unity unless B is $\mathfrak{p}$-adic and $f(\theta)=\mathfrak{p}$, in which case $R(\theta)$ is the ratio between a classical Gauss sum and its absolute value.

To simplify the statement of the arithmetic results: If A is a finite overfield of the local number field, B, then a character, θ, of B^* (i. e. a multiplicative character of B) is said to divide a character, Θ, of A^* (written $\theta|\Theta$) if Θ is obtained from θ by composition with the relative norm, $N_{A/B}$.

THEOREM 4. *Let B be a local number field,*

(a) *If A is a cyclic overfield of prime degree and Θ is a character of A^* divisible by a character of B^* then*

$$[\prod_{\theta|\Theta} R(\theta)]/R(\Theta)=V(A/B) \tag{10}$$

$$[\prod_{\theta|\Theta} \theta(-1)]/\Theta(-1)=[V(A/B)]^2 \tag{11}$$

the products being over all characters of B^* which divide Θ, $V(A/B)$ being a fourth root of unity which is independent of Θ. Furthermore $V(A/B)=1$ if degree A/B is odd.

(b) *If A is an abelian overfield of B with Galois group of type (p,p) and if θ_j $(j=1,2)$ is a character of A_j^*, A_j being an intermediate field of*

degree p over B, such that θ_1 and θ_2 divide the same character of A^*, neither being divisible by any character of B^* then

$$R(\theta_1)V(A_1/B) = R(\theta_2)V(A_2/B) \tag{12}$$

$$\theta_1(-1)[R(\theta_1)]^2 = \theta_2(-1)[R(\theta_2)]^2. \tag{13}$$

COROLLARY. *Equations* (10) *and* (11) *are also valid if A is any abelian extension of B.*[7] *If C is an immediate field then*

$$V(A/B) = V(A/C)[V(C/B)]^{\text{degree } A/C}. \tag{14}$$

The corollary is derived from the theorem by a simple induction argument. The theorem is proven by the arithmetic characterization of local abelian root numbers [4], [12]. As noted in the introduction, the proof is too long to be included in this paper.

If X is a linear character of a global Galois group, $G(K/k)$, then the global abelian root number, $W(X)$, may be written as a product of local abelian root numbers, [1, 2]:

$$W(X) = \prod_{\mathfrak{p}} R(X_{\mathfrak{p}}), \tag{15}$$

the product being over all primes of k. It should be noted that (15) need not be the only possible factorization of $W(X)$, further possibilities being covered by the following definition.

Definition 6. A function, H, defined on all multiplicative characters of local number fields is said to be a *factorization of unity* if it satisfies hypothesis (b) of Theorem 3 and if in that notation

$$1 = \prod_{\mathfrak{p}} H(X_{\mathfrak{p}}), \qquad (X \text{ linear}) \tag{16}$$

the product being over all primes of the ground field.

Theorem 3 and the pertinent parts of Theorem 1 may be reformulated so as to make Theorem 4 more directly applicable.

THEOREM 5. *Let F be a function defined on all multiplicative characters of local number fields such that*

(a) *In the notation of Theorem* 4a, $\prod_{\theta|\Theta} F(\theta) = F(\Theta)$.

(b) *In the notation of Theorem* 4b, $F(\theta_1) = F(\theta_2)$.

(c) *F satisfies conditions* (b) *and* (c) *of Theorem* 3.

---

[7] Theorem 4a in this more general form has been independently stated and partially verified by Hasse [5]. For a full proof see the dissertation referred to in footnote 1.

Then F is extendable with respect to the family of all local Galois groups. If $\bar{M}$ is the extension of M and $\bar{F}$ is the extension of F then for each character, X, of a global Galois group, $G(K/k)$,

$$\bar{M}(X) = \prod_{\mathfrak{p}} \bar{F}(X_{\mathfrak{p}}),$$

the product being over all primes of k and M being as in the notation of condition (c) *of Theorem* 3.

Proof. Let Δ be a family of nilpotent local Galois groups. We assert that F satisfies the conditions of Theorem 1 for Δ. F certainly satisfies conditions (a) and (d) of Theorem 1 and in the statement of condition (c) of that theorem $G = G_1$ and $G_0 = H$ as G is nilpotent. Hence condition (c) of Theorem 1 may for this application be written: If Θ is a linear character of H which is the restriction to H of a linear character of G then $F(\Theta) = \prod_{\theta|H=\Theta} F(\theta)$, the product being over all linear characters of G whose restriction to H is Θ. Furthermore in the statement of condition (b) of Theorem 1, as θ_i $(i = 1, 2)$ induces an irreducible character of G, it follows from the Frobenius reciprocity theorem that θ_i cannot be the restriction to G_i of a linear character of G. Also, as θ_1 and θ_2 induce the same irreducible character of G, it follows from Mackey, [8], that there exists $\sigma \varepsilon G$ such that $\theta_1{}^{\sigma}$ coincides with θ_2 on H. As $G = G_1G_2$ and H contains the commutator subgroup of G, it follows that θ_1 and θ_2 coincide on H. That hypothesis (a) of Theorem 5 implies that F satisfies condition (c) of Theorem 1 and that hypothesis (b) of Theorem 5 implies that F satisfies condition (b) of Theorem 1 now follows from a well known fact: Let B be a local number field, A a normal extension field of finite degree and C an intermediate field. If θ is a character of B^* which corresponds to a linear character, X, of $G(A/B)$ and θ_0 is the character of C^* which corresponds to the restriction of X to $G(A/C)$ then $\theta_0 = \theta \circ N_{C/B}$, the composition of θ with the relative norm of C over B.

It now follows from Theorem 1 that condition (a) of Theorem 3 is satisfied. This proves the extendability of F. The final assertion concerning the functions, $\bar{F}$ and $\bar{M}$ is a direct consequence of the uniqueness of $\bar{M}$ and the previously used argument of Artin [11], pp. 3-5.

Our main result may now be stated. It is a direct consequence of Theorems 4 and 5 and the fact that $\theta \to \theta(-1)$ is a factorization of unity.

THEOREM 5′. (a) *Let F be a function defined on linear characters of local Galois groups by setting $F(\theta) = [R(\theta)]^2\theta(-1)$ for each multiplicative character, θ, of a local number field* (*with the usual identification*), *then F*

is extendable with respect to the family of all local Galois groups. Denoting the extended function by F, $[W(X)]^2 = \prod_{\mathfrak{p}} F(X_{\mathfrak{p}})$, for each character, X, of a global Galois group, the product being over all primes of the ground field.

(b) *If H is a factorization of unity such that in the notation of*

$$\text{Theorem 4(a)},\ [\prod_{\theta|\Theta} H(\theta)]/(\Theta) = V(A/B)$$

$$\text{Theorem 4(b)},\ H(\theta_1)V(A_1/B) = H(\theta_2)V(A_2/B)$$

then $F'\colon \theta \to R(\theta)/H(\theta)$ defines a function on linear characters of local Galois groups which is extendable with respect to every family of local Galois groups and the extended function gives a factorization of the Artin root number in the obvious way.

4. The question of sign. The existence of a function, H, satisfying the conditions of Theorem 5(b) is still an open question and for this reason the existence of local non-abelian root numbers and the factorization of the Artin root number is established only to within factors ± 1. This aspect of the problem shall now be considered. As $V(A/B) = 1$ if degree A/B is odd (Theorem 4), it is to be expected that it is enough to determine $H(\theta)$ for θ of period 2^m (as used in this section 2^m is to be understood to be a generic integral power of 2). More precisely:

Lemma 4. *Let J be a function defined on multiplicative characters of period 2^m of local number fields such that*

(a) *If X is a linear character of period 2^m of a global Galois group, $G(K/k)$, then $J(X_{\mathfrak{p}}) = 1$ for almost all primes of k, and $\prod_{\mathfrak{p}} J(X_{\mathfrak{p}}) = 1$, the product being over all primes of k.*

(b) *In the notation of Theorem 4(a), if $\theta|\Theta$, θ of period 2^m and $q =$ degree A/B then*

(b_1) *$[J(\theta)]^q = J(\Theta)$ if $q \neq 2$*

(b_2) *$J(\theta)J(\theta\delta)/J(\Theta) = R(\delta)$ if $q = 2$ and δ is the character of B^* which "cuts out" A over B.*

(c) *In the notation of Theorem 4(b), if θ_1 and θ_2 are of period 2^m and degree $A/B = 4$ then $J(\theta_1)V(A_1/B) = J(\theta_2)V(A_2/B)$*

and if for each multiplicative character, θ, of a local number field, $\theta = \theta'\theta''$ denotes the natural decomposition of θ into a character, θ', of odd period and a character, θ'', of period 2^m then $H\colon \theta \to J(\theta'')$ is a factorization of unity which satisfies the conditions of Theorem 5(b).

Proof. It is enough to show that H satisfies the condition corresponding to Theorem 4(b). In the statement of that theorem, let $p^2 =$ degree A/B, and let $\theta_i = \theta_{i1}\theta_{i2}$, $(i = 1, 2)$ be the natural decomposition of θ_i into a character, θ_{i1}, of odd period and a character, θ_{i2}, of period 2^m. It is enough to show that

$$J(\theta_{12})V(A_1/B) = J(\theta_{22})V(A_2/B). \tag{17}$$

Let σ be a non-trivial element of the Galois group, $G(A/A_2) \cong G(A_1/B)$, then since $\theta_1 \circ N_{A/A_1} = \theta_2 \circ N_{A/A_2}$, it follows that $\theta_1^{\sigma-1} \circ N_{A/A_1}$ is the principal character of A^*. Hence $\theta_1^{\sigma-1}$ is trivial on $N_{A/A_1}A^*$ but is not trivial on A_1^* as otherwise (since A_1 is cyclic over B) θ_1 is divisible by a character of B^*, contrary to hypothesis. Hence $\theta_1^{\sigma-1}$ is of period p. Thus $\theta_{12}^{\sigma-1}$ is trivial if and only if $p \neq 2$. It is clear that θ_{11} and θ_{21} divide the same character of A^* and that the same holds for characters, θ_{12} and θ_{22}. If p is odd then θ_{12} is divisible by a character, θ, of period 2^m, of B^* and by an easy argument $\theta | \theta_{22}$, whence (17) follows from (b_1). If $p = 2$ then θ_{12} is not divisible by any character of B^* so that (17) follows from (c).

A partial solution for J may be given.

LEMMA 5. *If θ is a real multiplicative character of a local number field let $J(\theta) = R(\theta)$. J satisfies the conditions of the last lemma for characters of this type.*

Proof. If X is a linear character of period 2^m of a global Galois group, $G(K/k)$, such that $X_{\mathfrak{p}}$ is real for each prime $\mathfrak{p}$, of k, then X is of period 2 or 1. In either case $1 = W(X) = \prod_{\mathfrak{p}} R(X_{\mathfrak{p}})$, so that condition (a) of the last lemma is satisfied. Condition (b_2) and (c) follow from Theorem 4, and for condition (b_1) let C be the quadratic extension of B cut out by θ, then AC is the extension of A cut out by Θ and AC is an abelian extension of B. Hence $R(\Theta) = V(AC/A)$, $R(\theta) = V(C/B)$ and by (14),

$$V(AC/A)[V(A/B)]^2 = V(AC/B) = V(AC/C)[V(C/B)]^q,$$

whence (c_1) follows from $V(A/B) = V(AC/C) = 1$.

It follows from this lemma that if Y is a character of a local Galois group, G, then the (local) root number of Y is well defined if no cyclic group of order 4 is in the family generated by G. Furthermore if X is a character of a global Galois group, G, then the Artin root number, $W(X)$, has a well defined decomposition if the decomposition subgroups of G satisfy this condition.

Theorem 5′ indicates that the determination of H requires a canonical

extraction of the square root of $\theta(-1)$. If θ is of period 2 then this is provided by $R(\theta)$.

We shall now show that in a certain sense it is impossible to find a function H which satisfies the conditions of Theorem 5′b. If H were found then for each multiplicative character, θ, of a local number field, B, $R(\theta)/H(\theta)$ could be viewed as a new local abelian root number. It would therefore be expected that H would satisfy the further conditions:

(I) $H(\theta)=1$ if θ is the principal character of B^*

(II) $H(\theta)=1$ if θ is an absolutely unramified character (i.e. if the kernel of θ determines a cyclic extension of B which is absolutely unramified).

It will be shown that these conditions are incompatible with the previous conditions on H. Suppose that there exists an H satisfying all the conditions. It follows from (I) that if θ is real then $H(\theta)=R(\theta)$. In particular, if B is the field of real numbers then $1/H(\theta)=\sqrt{[\theta(-1)]}$, where $\sqrt{1}=1$, $\sqrt{(-1)}=i$. If n is any positive integer, let G_n be the Galois group of $Q(\zeta_n)$ over Q, where ζ_n is a primitive n-th root of unity. If p is an odd rational prime number and X is a character of G_p then (from II), $H(X_\infty)H(X_p)=1$. Likewise if q is an odd prime and Y is a character of G_q then $H(Y_\infty)H(Y_q)=1$, while if XY is the character of G_{pq}, $(p\neq q)$, obtained in the obvious way then $H(X_\infty Y_\infty)H(X_pY_p)H(X_qY_q)=1$. As X_q is unramified and $X_q(q)X_p(q)=1$, X_q is trivial if $X_p(q)=1$, which is certainly true if $q\equiv 1 \pmod p$. If q satisfies this condition then

$$H(X_pY_p)=\sqrt{[X_p(-1)Y_q(-1)]}/\sqrt{[Y_q(-1)]}.$$

Given p, X may be chosen so that $X_p(-1)=-1$, in which case $H(X_pY_p)=\sqrt{[-Y_q(-1)]}/\sqrt{[Y_q(-1)]}$. Hence to obtain a contradiction, it is enough to choose p and an unramified character θ of $Q_p{}^*$ such that $Y_q(-1)$ is not independent of the choice of q and Y which satisfy the conditions: $q\equiv 1 \pmod p$, $Y_p=\theta$ (ie. $1/\theta(p)=Y_q(p)$). To do this let $p=5$, $\theta(5)=i$. For $q=41$, p is a quadratic but not a biquadratic residue mod q, likewise for $q=241$. Hence in either case there exists Y of period 8 such that $Y_q(p)=-i$. As -1 is an eighth power mod 241 but not mod 41, $Y_q(-1)=-1$ for $q=41$, $Y_q(-1)=1$ for $q=241$. The contradiction is thus proven.

This counter example almost assures us of the nonexistence of a function H satisfying the conditions of Theorem 5′b. The problem of finding H is that of reformulating the definition of local abelian root numbers so that

the group theoretical properties of the Artin root number may be explained locally. Instead we can consider the global relations between local abelian root numbers which can be deduced from the properties of the Artin root number. Let K be a normal overfield of an algebraic number field, k; let ϕ_i $(i=1,\cdots,r)$ be a linear character of a subgroup, $G(K/k_i)$ of $G(K/k)$ such that $\sum_{i=1}^{r} a_i X_i = 0$, where X_i is the character of $G(K/k)$ which is induced by ϕ_i and a_i is an integer. The global relations referred to above are of the form

$$\prod_{\mathfrak{p}} \prod_{i=1}^{r} \prod_{\mathfrak{q}_i | \mathfrak{p}} R((\phi_i)_{\mathfrak{q}_i})^{a_i} = 1$$

($\mathfrak{p}$ runs through all primes of k; $\mathfrak{q}_i$ runs through all primes of k_i). Following Hasse, let δ_i be the character of $G(K/k)$ induced by the principal character of $G(K/k_i)$. Let $|\delta_i|$ denote the linear character of $G(K/k)$ obtained by taking the determinant of the matrix representation corresponding to δ_i. As a direct consequence of Theorem 5′a, the $\mathfrak{p}$ constituent of the above relation is given by

$$\prod_{i=1}^{r} \prod_{\mathfrak{q}_i | \mathfrak{p}} R((\phi_i)_{\mathfrak{q}_i})^{a_i} = \pm \prod_{i=1}^{r} [R(|\delta_i|_{\mathfrak{p}})]^{-a_i}, \qquad \text{(cf. [5, page 83])}$$

where the inner product on the left is over all primes of k_i which divide $\mathfrak{p}$. As $|\delta_i|$ is either the principal character or a quadratic character of $G(K/k)$, it is natural to conjecture that all of these global relations are products of two kinds of relations:

1. If X is any linear character of $G(K/k)$ then $\prod_{\mathfrak{p}} X_{\mathfrak{p}}(-1) = 1$.

2. If X is a quadratic character of $G(K/k)$ then $\prod_{\mathfrak{p}} R(X_{\mathfrak{p}}) = 1$.

(The product in both cases is over all primes of k.)

5. Local root numbers. In this section let C^* be the multiplicative group of non-zero complex numbers and let Γ be the factor group, $C^*/\{\pm 1\}$. If X is a character of a local Galois group, $G = G(K/k)$, let $W(X)$ denote the *local root number* $(\varepsilon\,\Gamma)$ which is obtained by extending the function, $X \rightarrow R(X)\sqrt{[X(-1)]}$, defined on linear characters of local Galois groups (Theorem 5′(a)). It follows from Definition 5 that if k is Archimedean then $W(X) = 1$, the identity element of Γ. Hence in studying local root numbers it may be assumed that k is a $\mathfrak{p}$-adic number field. In the $\mathfrak{p}$-adic case we define (following Hasse) the *local Galois Gauss* sum, $\tau(X) = W(X)\sqrt{[Nf(X)]}$,

where $Nf(X)$ is the absolute norm of the conductor of X. If Y is a character of a global Galois group, $G(E/F)$, then Hasse's Galois Gauss sum, $\tau(Y)$, is defined by

$$\tau(Y) = W(Y)\sqrt{[Nf(Y)]}\,\varepsilon\prod_{\mathfrak{p}}\tau(Y_{\mathfrak{p}}), \tag{18}$$

the product over all primes of F being understood to be a coset of $\{\pm 1\}$. In the following the symbols, $W(X)$, $\tau(X)$, shall refer to elements of C^* if X is a character of a global Galois group and elements of Γ (or cosets of $\{\pm 1\}$) if X is a character of a local Galois group. An element of Γ will be said to be a root of unity (resp.: an algebraic integer, resp.: an element of a certain algebraic number field) if the elements of the corresponding coset of $\{\pm 1\}$ have this property.

Continuing with the discussion of the local situation, let $\mathfrak{W}_0$ be the inertial subgroup, $\mathfrak{W}_1, \mathfrak{W}_2, \cdots$ the ramification subgroups of G. Let

$$m(X) = [\mathfrak{W}_0]^{-1}\sum_{i=0}^{s}[\mathfrak{W}_i], \tag{19}$$

where s is the smallest integer such that X is trivial (i. e. identically $X(1)$) on $\mathfrak{W}_{s+1}$, $m(X)$ being understood to be zero if $s = -1$. Letting $\mathfrak{p}$ be the prime ideal of k, the conductor, $f(X)$, is given by Artin [11],

$$\operatorname{ord}_{\mathfrak{p}} f(X) = [\mathfrak{W}_0]^{-1}\sum_{i=0}^{\infty}\{[\mathfrak{W}_i]X(1) - X(\mathfrak{W}_i)\}, \tag{20}$$

where $X(\mathfrak{W}_i) = \sum_{w\in\mathfrak{W}_i} X(w)$, this last sum being over all elements of $\mathfrak{W}_i$.

LEMMA 6. *If X is an irreducible character of G then*

$$m(X) = [\operatorname{ord}_{\mathfrak{p}} f(X)]/X(1) \tag{21}$$

and if $X = \sum_{j=1}^{r} X_j$, a sum of characters of G then

$$m(X) = \operatorname{Max}_{1\leq j\leq r} m(X_j). \tag{22}$$

Proof. From the orthogonality relations, $n_i = [\mathfrak{W}_i]^{-1}X(\mathfrak{W}_i)$ is the number of times the trivial character of $\mathfrak{W}_i$ occurs in $X|\mathfrak{W}_i$. If X is irreducible then $X|\mathfrak{W}_i$ is a sum of conjugate characters of $\mathfrak{W}_i$ as $\mathfrak{W}_i$ is an invariant subgroup of G. Hence n_i is either $X(1)$ or zero depending upon whether or not i exceeds s. Thus (21) follows from (19) and (20). If $X = \sum X_j$ and X is trivial on $\mathfrak{W}_i$ then each X_j is trivial on that group, i. e. $m(X) \geqq m(X_j)$. Conversely if each X_j $(j = 1, \cdots, r)$ is trivial on $\mathfrak{W}_i$ then X is trivial on $\mathfrak{W}_i$, i. e. $\operatorname{Max}_j m(X_j) \geqq m(X)$ which proves (22).

It follows from the definition that either $m(X) = 0$ or $m(X) \geqq 1$. It

is easily shown by the methods of Artin, [11, page 9], that $m(X)$ is a rational number whose denominator is a power of the rational prime, p, which is divisible by $\mathfrak{p}$. $m(X)$ need not be an integer, a counter example being provided by the situation discussed in Theorem 4(b). In the notation of that theorem let C be the cyclic extension of A which is cut out by the character, $\theta_1 \circ N_{A/A_1}$, of A^*. Let A be totally and wildly ramified over B and let θ_1 be chosen so as to have the "minimal" conductor (i. e. with smallest possible exponent) and let X be the (irreducible) character of $G(C/B)$ which is induced by the linear character of $G(C/A_1)$ which corresponds to θ_1. Under these circumstances $m(X)$ is not an integer but the details of the computation are not needed at this time.

We now show that if X is irreducible then $m(X)$ determines the nature of the local root number, $W(X)$.

THEOREM 6. *Let X be an irreducible character of the local Galois group, $G(K/k)$.*

(a) *If $m(X) \leqq 1$ then there exists an intermediate field, k', unramified over k such that X is induced by a linear character, Z, of $G(K/k')$ of conductor $\mathfrak{q}^{m(X)}$ ($\mathfrak{q}$ = prime of k') and $\tau(X) = \tau(Z)$, hence an algebraic integer.*

(b) *If $m(X) \neq 1$ then $W(X)$ is a root of unity.*

Proof. If $m(X) = 0$ then X is trivial on $\mathfrak{W}_0$, whence X is linear and the theorem is trivial in that case. If $m(X) = 1$ then X is trivial on $\mathfrak{W}_1$, whence it is induced by a linear character, Z, of a subgroup, G', of G which contains $\mathfrak{W}_0$ (as $\mathfrak{W}_0/\mathfrak{W}_1$ is cyclic, hence certainly abelian). G' corresponds to a subfield, k', of K which is unramified over k, so that $N_{k'/k}f(Z) = f(X)$, whence $m(Z) = 1$. Clearly $Nf(X) = Nf(Z)$ so that $\tau(X) = \tau(Z)$ which completes the proof of (a).

If $m(X) > 1$ then there exists a minimal integer, s, $s > 0$, such that X is trivial on $\mathfrak{W}_{s+1}$. As $\mathfrak{W}_s/\mathfrak{W}_{s+1}$ is abelian, X is induced by a character, X', of a subgroup, G', of $G(K/k)$ which contains $\mathfrak{W}_s$ and leaves δ fixed, δ being one of the linear characters of $\mathfrak{W}_s$ which occurs in $X|\mathfrak{W}_s$ (Lemma 2). For the same reason, X' has a Brauer decomposition in characters of G' which are induced by linear characters, Y_j, of subgroups, G_j, of G' which contain $\mathfrak{W}_s$, such that $\delta = Y_j|\mathfrak{W}_s$. δ is a non-trivial character of $\mathfrak{W}_s$ as $X|\mathfrak{W}_s$ contains only conjugates of δ, hence the conductor of δ is $\mathfrak{P}^r$, $r > 1$, $\mathfrak{P}$ being the prime of V_s, the s-th ramification subfield of K over k. If k_j is the subfield of K which corresponds to G_j then $V_s \supset k_j \supset k$ and Y_j corresponds to a linear character, θ_j, of k_j^*. Also $\theta_j \circ N_{V_s/k_j} = \Delta$, the character of V_s^* which corresponds to δ. As $N_{V_s/k_j}(1 + \mathfrak{P}) \subset 1 + \mathfrak{p}_j$ ($\mathfrak{p}_j$ = prime of k_j), if $\mathfrak{p}_j^2 \nmid f(\theta_j)$

then $\mathfrak{P}^2 \nmid f(\Delta) = f(\delta)$, a contradiction. Hence $\mathfrak{p}_j{}^2 | f(\theta_j)$ so that the abelian root number, $R(\theta_j)$, is a root of unity. As $W(X)$ is a product of such quantities and their inverses, and powers of i, (b) follows from the previous discussion of the case $m(X) = 0$.

As an immediate consequence of this theorem and equation (18) we have a conjecture of Hasse:

COROLLARY. *If X is a character of a global Galois group then the Galois Gauss sum, $\tau(X)$, is an algebraic integer.*

We shall now consider Hasse's conjecture concerning the algebraic number field in which the Artin root number lies. It follows from the definitions that if X is a linear character of a local Galois group, $G(K/k)$, k being a $\mathfrak{p}$-adic number field, then $[W(X)]^2 = [R(X)]^2 X(-1)$ lies in the field, A, obtained by adjoining to Q (= the field of rational numbers) the values assumed by X and a primitive N-th root of unity, where N is the power of p (= the rational prime divisible by $\mathfrak{p}$) defined by the ideal theoretic relation

$$1/N = S_{k/Q_p}(1/(\mathfrak{D}_k f(X))), \tag{23}$$

Q_p being the p-adic completion of Q. $W(X)$ itself lies in the field obtained by adjoining to A the elements i and $\sqrt{p}$. The generalization of this result is complicated by the uncertain status of the question of sign (Section 4).

THEOREM 7. *If X is a character of a local Galois group, $G(K/k)$, k being a $\mathfrak{p}$-adic number field, then $[W(X)]^2$ lies in the field $A = Q(X; m(X)/e)$ obtained by adjoining to Q the values assumed by X and the N roots of unity, where $N = p^r$, p is the rational prime divisible by $\mathfrak{p}$, r is the smallest integer not less than $m(X)/e$, e being the absolute ramification of k.*

Proof. The theorem is first proven for X irreducible. If X is linear then the assertion follows from the fact that $S_{k/Q_p}(1/(\mathfrak{D}_k f(X))) = p^{[-m(X)/e]}$, the brackets denoting the largest integer not greater than $-m(X)/e$. If $m(X) = 1$ then by Theorem 6, $W(X) = W(Z)$, Z being a linear character of $G(K/k')$ where k' is an unramified extension of k. As $m(Z) = 1$ the assertion follows from the discussion of the linear case. If $m(X) > 1$ then in the notation of the proof of the preceding theorem, $[W(X)]^2$ lies in the composition of fields of the type $Q(Y_j; m(Y_j)/e_j)$, Y_j being a linear character of $G_j = G(K/k_j)$ which contains $\mathfrak{V}_s$, e_j being the absolute ramification of k_j. Furthermore $Y_j | \mathfrak{V}_s = \delta$. As the i-th ramification subgroup of G_j is $\mathfrak{V}_i \cap G_j$,

$\mathfrak{W}_s$ is the s-th ramification subgroup of G_j and $\mathfrak{W}_{s+1}$ is the $(s+1)$-th ramification subgroup of G_j. Y_j is trivial on $\mathfrak{W}_{s+1}$ but not on $\mathfrak{W}_s$. Hence

$$m(Y_j)/e_j = (e_j[\mathfrak{W}_0 \cap G_j])^{-1} \sum_{i=0}^{s} [G_j \cap \mathfrak{W}_i]. \tag{24}$$

But $e_j[\mathfrak{W}_0 \cap G_j] =$ absolute ramification of $K = e[\mathfrak{W}_0]$ so that $m(Y_j)e_j \leqq m(X)/e$. It follows that if $Y_1, \cdots, Y_t$ are the linear characters of subgroups of G which are involved in the Brauer decomposition of X described in the proof of Theorem 6 then $[W(X)]^2 \varepsilon\, Q(Y_1, \cdots, Y_t; m(X)/e) = B$, it being understood that the field, B, contains all values assumed by $Y_1, \cdots, Y_t$ respectively. Certainly $B \supset A$, but A contains all roots of unity introduced by the additive characters of the fields $k_1, \cdots, k_t$ in forming the abelian Gauss sums associated with the linear characters $Y_1, \cdots, Y_t$. The Brauer decomposition of X may be written

$$X = \sum_{j=1}^{t} a_j X_j, \tag{25}$$

X_j being the character of G which is induced by Y_j, and each a_j being an integer. Clearly, $W(X) = \prod_{j=1}^{t} W(Y_j)^{a_j}$. If σ is an isomorphism of B which leaves the elements of A fixed then it leaves each value of X fixed so that (25) may be written with each X_j replaced by the character $\sigma \circ X_j$ which is induced by $Y_j^\sigma = \sigma \circ Y_j$. Hence $W(X) = \prod_{j=1}^{t} W(Y_j^\sigma)^{a_j}$. But σ leaves fixed the values assumed by the additive characters, i.e. $[W(Y_j^\sigma)]^2 = [W(Y_j)^\sigma]^2$. It follows that σ leaves $[W(X)]^2$ invariant, whence $[W(X)]^2$ lies in A.

If $X = \sum_i L_i$, a sum of irreducible characters of G then by the preceding argument and (22) it follows that $[W(X)]^2 \subset Q(L_1, \cdots, L_r, m(X)/e) = A'$. The same argument as in the irreducible case shows that if σ is an isomorphism of A' which leaves the elements of A fixed then

$$[W(X)]^2 = [\prod_i W(L_i^\sigma)]^2 = [W(X)^\sigma]^2$$

as σ leaves invariant all values assumed by additive characters used in expressing $W(L_i)$ in terms of local abelian root numbers.

This result has an obvious global consequence.

COROLLARY 1. *If X is a character of a global Galois group, $G(K/k)$, then $[W(X)]^2$ lies in the field A obtained by adjoining to Q the values assumed by X and a primitive N-th root of unity, where*

$$\log N = \sum_p (\log p) \max_{\mathfrak{p} | p} m_{\mathfrak{p}}, \tag{27}$$

$m_{\mathfrak{p}}$ being the smallest integer which is not less than $m(X_{\mathfrak{p}})/e_{\mathfrak{p}}$, $e_{\mathfrak{p}}$ being the absolute ramification of $\mathfrak{p}$, the sum being over all finite rational primes.

The following corollaries refer to both local and global characters. The notation is as in Theorem 7 in the local case and as in Corollary 1 in the global case. If n is an integer let Z_n denote a primitive n-th root of unity.

COROLLARY 2. *$W(X)$ lies in $Q(Z_r)$, where r is the least common multiple of 8, N and the degree, n, of K over k.*

Proof. In the local case it follows by the same reasoning as in Theorem 7 (without the automorphism argument as the values assumed by the characters, Y_j, are n-th roots of unity) that $W(X) \varepsilon Q(Z_n, Z_N, i, \sqrt{p})$. If $m(X) = 0$ then $W(X)$ is an n-th root of unity, while if $m(X) > 0$ then $Z_p \varepsilon Q(Z_N)$. As i, $\sqrt{p}$, lie in $Q(Z_p, i)$ for $p \neq 2$, and i, $\sqrt{2}$ lie in $Q(Z_8)$ the assertion follows in the local case. The proof in the global case is now immediate.

COROLLARY 3. *If the question of sign is solved for the local Galois groups involved with $G(K/k)$ (in particular if the family generated by $G(K/k)$ contains no cyclic group of order 4) then $W(X) \varepsilon A(Z_8)$.*

The proof is a repetition of previous arguments. We note that everything said here for root numbers (i. e. in Theorem 7 and its corollaries) applies equally well to Galois Gauss sums.

Final Remarks. The proof of the integrality of Galois Gauss sums requires only a weak form of Theorem 5′. It is enough to know the extendability of the local abelian root number considered as a function which takes its values in the group C^*/E, where E is the group of all roots of unity in the complex field. It follows that a weak form of Theorem 4 is adequate, it being enough to know that (10) and (12) are valid as congruences modulo E. These congruences are a trivial consequence of the multiplicative identities of [12] and in fact as a congruence, (12) is a direct consequence of (10).

The results of this paper may be extended without difficulty to the root number in the functional equation of the L-series studied by Weil in [14]. This may be done most easily by using the fact that an irreducible character of one of Weil's decomposition groups is the product of a linear unramified character and an irreducible character with kernel of finite index and the latter may be identified with a character of a local Galois group. While the root numbers appearing in this theory need not be algebraic numbers, Theorem 6 may be rewritten so as to remain valid.

Finally it may be noted that the methods of this paper may be used to prove the existence and most of the properties of the Artin conductor, [11], without the use of Artin's specific formula.

HARVARD UNIVERSITY.

REFERENCES.

[1] E. Artin, "Zur Theorie der L-Reihen mit allgemeinen Gruppencharakteren," *Abhandlungen aus dem Mathematischen Seminar der Universität Hamburg*, vol. 8 (1931), pp. 292-306.

[2] J. Tate, "Fourier Analysis in Number Fields and Hecke's Zeta Functions," Thesis, Princeton University, 1950 (unpublished).

[3] H. Hasse, "Allgemeine Theorie der Gaussche Summen in algebraischen Zahlkorpern," *Abhandlungen der Deutschen Akademie der Wissenschaften zu Berlin*, Math-natur-wiss. Kl. Jahrg. 1951, Nr. 1, pp. 1-23.

[4] E. Lamprecht, "Allgemeine Theorie der Gaussche Summen in endlichen kommutativen Ringen," *Mathematische Nachrichten*, vol. 9 (1935), pp. 149-196.

[5] H. Hasse, "Artinsche Führer, Artinsche L-Funktionen und Gaussche Summen uber Endlich-Algebraischen Zahlkorpern," *Acta Salmanticensia*, Ciencias: Seccion de Mathematicas, 1954.

[6] R. Brauer, "On Artin's L-series with general group characters," *Annals of Mathematics*, vol. 48 (1947), pp. 502-514.

[7] E. Artin, "Galois Theory" (second edition), Notre Dame Mathematical Lectures, Notre Dame, Indiana (1946).

[8] G. Mackey, "On induced representations of groups," *American Journal of Mathematics*, vol. 73 (1951), pp. 576-592.

[9] K. Taketa, "Uber die Gruppen deren Darstellung samtlich auf monomiale gestalt transformieren lassen," *Proceedings of the Imperial Academy of Japan*, vol. 6 (1930).

[10] E. Artin, "Algebraic numbers and algebraic functions," vol. I, Princeton University, 1950-1951.

[11] ———, "Die gruppentheoretische Struktur der Diskriminanten algebraischer Zahlkörper," *Journal für die Reine und Angewandte Mathematik*, vol. 164 (1931), pp. 1-11.

[12] H. Davenport and H. Hasse, "Die Nullstellen der Kongruenzzetafunktionen in gewissen zyklischen Fällen," *Journal für die Reine und Angewandte Mathematik*, vol. 172 (1934), pp. 151-182.

[13] B. Huppert, "Normalteiler und maximale Untergruppen endlicher Gruppen," *Mathematische Zeitschrift*, vol. 60 (1954), pp. 409-434.

[14] A. Weil, "Sur la théorie du corps de classes," *Journal of the Mathemical Society of Japan*, vol. 3 (1951), pp. 1-35.

to
],

n,"
rg,

is,

hl-
zu

au-
.
hen
as:

he-

es,

he-

alt
an,

ni-

her
164

in
he-

n,"

ety